University of Alberta

Investigation of a putative RNA helicase in the developmental program of *Streptomyces coelicolor*

by

Kent David Gislason ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

in

Microbiology and Biotechnology

Department of Biological Sciences

Edmonton, Alberta

Spring, 2005

0-494-08065-5

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN:
Our file *Notre reference*
ISBN:

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

S. coelicolor embodies a unique developmental program. *sco3550*, which encodes a putative RNA helicase, was originally identified due to its close genomic proximity to the developmental gene, *bldG*. Western analysis of SCO3550 revealed a doublet corresponding to the predicted molecular weight of native SCO3550. In order to analyze the phenotype of a *sco3550* disruption mutant, PCR-targeted REDIRECT© mutagenesis was attempted. However, *sco3550* appears resistant to complete disruption and putative *sco3550* disruption strains revealed a mixture of wild-type and mutant chromosomes. To confirm the biochemical nature of SCO3550 as a RNA helicase, *in vitro* RNA unwinding assays were performed. The results of the *in vitro* assays suggest that SCO3550 may act as a specific RNA binding protein to effect dsRNA destabilization *in vivo*. To examine potential protein-protein interactions with SCO3550, a HIS_6-SCO3550 recombinant protein was overexpressed in *S. coelicolor* and chemical crosslinking studies revealed the presence of a potential high MW complex.

Acknowledgements

This thesis project has by no means been a solitary endeavor and I would like to extend my gratitude to my supervisor Dr. Brenda Leskiw for a wealth of personal and professional mentorship over the past three years. I have gained so many valuable skills throughout this degree and my learning has been allowed to grow in no small part due to Brenda's generous assistance. I would also like to thank my committee members, Dr. George Owttrim and Dr. Susan Jensen, for many helpful discussions and for taking a genuine interest in my work throughout the course of this study.

The Leskiw lab has been a fantastic work and learning environment and I am indebted to the many members of the lab, past and present, who have been steadfast in their support. It's difficult to imagine what state my thesis project would be in today if not for the help and support from my fellow lab members: Dawn Bignell, Julie Stoehr, Claire Galibois, Kim Colvin, Linda Bui, Archana Parashar and Wisdom Ayidzoe.

Finally, I must thank my family and friends, especially my parents, my brother, Connor, and my girlfriend, Audrey, who have always provided a haven for me.

Table of Contents

		Page
Chapter 1:	**Introduction**	**1**
	1.0 General Overview of the Streptomycetes	2
	1.1 Physiological Adaptations for Survival in a Diverse Soil Ecosystem	2
	1.2 Completion and Extensions of the *Streptomyces coelicolor* Genome Sequence	8
	1.3 Genetic Characterization of Morphological Differentiation	11
	1.4 *whi* Mutants	11
	1.5 *bld* Mutants- a Link to Stress, Metabolism and Development	19
	1.6 RNA Helicases	39
	1.7 The RNA Folding Problem	40
	1.8 RNA Helicase Taxonomy and the Structure and Function of Conserved Motifs	42
	1.9 The Action of RNA Helicases in Nature- a Broadening Definition	49
	2.0 Thesis Objectives	55
Chapter 2:	**Materials and Methods**	**58**
	2.1 Bacterial Strains, Plasmids and Growth Conditions	59
	2.1.1 *Streptomyces coelicolor* and *Escherichia coli* strains	59
	2.1.2 Plasmid vectors	59
	2.1.3 Growth and maintenance of *E. coli* strains	59
	2.1.4 Preparation of *E. coli* glycerol stocks	67
	2.1.5 Growth and maintenance of *S. coelicolor* strains	67
	2.1.6 Preparation of *S. coelicolor* glycerol stocks	67
	2.2 DNA Isolation and Transformation	68
	2.2.1 Preparation of *E. coli* competent cells	68

2.2.2 Transformation of *E. coli* 69
2.2.3 Isolation of plasmid DNA from *E. coli* 70
2.2.4 Preparation of *S. coelicolor* protoplasts..... 70
2.2.5 Transformation of *S. coelicolor* 70
2.2.6 *E. coli* conjugation into *S. coelicolor* 70
2.2.7 Isolation of chromosomal DNA from *S. coelicolor* 72

2.3 DNA Purification and Analysis..... 72
2.3.1 Restriction digestion and cloning of DNA..... 72
2.3.2 Polymerase Chain Reaction (PCR)..... 73
2.3.3 Agarose gel electrophoresis of DNA..... 77
2.3.4 DNA purification from agarose gels..... 78
2.3.5 Polyacrylamide gel electrophoresis of DNA..... 78
2.3.6 DNA purification from polyacrylamide gels..... 78
2.3.7 DNA Sequencing..... 79
2.3.8 Southern hybridization analysis..... 80
2.3.9 Colony hybridization analysis..... 80
2.3.10 Labeling of DNA probes..... 81
2.3.11 DNA hybridizations..... 83

2.4 RNA Analysis..... 84
2.4.1 Isolation of RNA..... 84
2.4.2 High resolution S1 nuclease mapping..... 86
2.4.3 Primer extension analysis..... 88

2.5 *E. coli* Protein Overexpression and Purification..... 89
2.5.1 Overexpression of MBP-SCO3550..... 89
2.5.2 Purification of MBP-SCO3550..... 90
2.5.3 Preparation of *S. coelicolor* cell-free extracts..... 91
2.5.4 Overexpression of SCO3550 in *S. coelicolor*..... 92
2.5.5 Purification of HIS_6-SCO3550 from *S. coelicolor* 93
2.5.6 Protein quantification..... 94
2.5.7 Protein Precipitation..... 94

2.6 Protein Analysis..... 95
2.6.1 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)..... 95
2.6.2 Preparation of poly-clonal antibodies..... 96
2.6.3 Western hybridization analysis..... 97
2.6.4 Affinity purification of poly-clonal antibodies..... 99

2.7 *in vitro* Biochemical Analysis..... 100
2.7.1 RNA helicase assays..... 100

2.8 Computer assisted sequence analysis..... 100

Chapter 3: Results..... 102

3.1 Sequence Analysis of *sco3550* 109

3.2 MBP-SCO3550 Overexpression and Purification in *E. coli* 110

3.3 Protein Analysis..... 120
3.3.1 Western analysis of SCO3550..... 120
3.3.2 Western analysis of SCO3550 using affinity-purified antibodies..... 121

3.4 Mutational Studies of *sco3550* 125
3.4.1 Deletion of *sco3550* in *S. coelicolor* 128
3.4.2 Construction of a conditional *sco3550* deletion mutant in *S. coelicolor* 150
3.4.3 Analysis of the *bldG* 3b/*sco3550* double mutant..... 154

3.5 Western Analysis Revisited..... 163
3.5.1 Western analysis with a positive control..... 163
3.5.2 Western analysis of SCO3550 in *bldG* 1DB mutant extracts. 169
3.5.3 Western analysis of SCO3550 in *S. coelicolor bldG* 1DB + pAU69 complemented extracts..... 173

3.6 Transcript Analysis of *sco3550* 174
3.6.1 S1 nuclease mapping of *sco3550* transcripts..... 175
3.6.2 Primer extension analysis of *sco3550* transcripts..... 182

3.7 Biochemical Characterization of Recombinant MBP-SCO3550..... 182
3.7.1 *in vitro* RNA unwinding assays..... 183

3.8 HIS_6-SCO3550 Overexpression and Binding Studies in *Streptomyces* 192
3.8.1 Protein overexpression and purification in *Streptomyces* 192
3.8.2 *in vivo* Chemical crosslinking experiments..... 200
3.8.3 *in vitro* Chemical crosslinking experiments..... 204

Chapter 4: Discussion..... 212

Chapter 5: References..... 227

List of Tables

Page

Chapter 2

2.1 *Escherichia coli* strains used in this study..... 60

2.2 *Streptomyces* strains used in this study..... 61

2.3 Plasmids and cloning vectors used in this study..... 62

2.4 Recombinant plasmids and cosmids constructed and used in this study..... 65

2.5 Oligonucleotide primers used in this study..... 75

List of Figures

Page

Chapter 1

1.1 General life cycle of the Streptomycetes..... 5

1.2 The Proposed *bld* gene extracellular signaling cascade..... 26

1.3 Structural organization of conserved helicase motifs..... 47

Chapter 3

3.1 Schematic representation of the *bldG* locus..... 105

3.2 Protein sequence alignment of SCO3550 with putative RNA helicase homologs found in other organisms..... 112

3.3 Schematic depiction of pAU330 construction..... 116

3.4 Overexpression and purification of MBP-SCO3550..... 119

3.5 Western analysis of SCO3550 protein by 7.5 % SDS-PAGE.... 123

3.6 Western analysis of SCO3550 using affinity-purified antibodies (6 % SDS-PAGE)..... 127

3.7 Schematic protocol for PCR-targeting of *sco3550* deletion in cosmid SCH5..... 131

3.8 Confirmation of *sco3550* gene replacement in pAU333..... 134

3.9 Schematic protocol for replacement of *sco3550* in the *S. coelicolor* chromosome..... 136

3.10 Verification of *sco3550* replacement in *S. coelicolor* M600... 140

3.11 Verification of *sco3550* deletion in the M145 chromosome..... 144

3.12 Verification of *sco3550* replacement in *S. coelicolor* M145.... 146

3.13 Phenotypic effect of SCO3550 overexpression from pAU332 in *S. coelicolor* 152

3.14 Conditional mutagenesis of *sco3550* in a *bldG* 3b/*sco3550* double mutant background..... 157

3.15 Western analysis of cell extracts from *S. coelicolor bldG* 3b/ *sco3550* double mutant strains..... 160

3.16 Western analysis revisited..... 166

3.17 DNA sequence of the *bldG-sco3550* intergenic region..... 168

3.18 Western analysis of SCO3550 in *bldG* 1DB mutant extracts... 172

3.19 5' S1 mapping of *sco3550* 177

3.20 Internal S1 nuclease protection analysis of *sco3550* transcripts 180

3.21 5'-3' unwinding activity of MBP-SCO3550..... 185

3.22 3'-5' unwinding activity of MBP-SCO3550..... 190

3.23 Western analysis of a plate culture induction of HIS_6-SCO3550 195

3.24 Liquid culture induction of HIS_6-SCO3550 expression and Ni^{2+}-NTA affinity chromotography..... 198

3.25 *in vivo* protein-protein crosslinking with DMS..... 203

3.26 *in vitro* protein-protein crosslinking with BS^3..... 207

3.27 Ni-NTA affinity chromotography of BS^3 crosslinked samples.. 210

List of Abbreviations

α	Alpha
β	Beta
γ	Gamma
λ	Lambda
σ	Sigma
A	Adenine (DNA) or Alanine
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
bp	Base pair
BS^3	bis(sulfosuccinimidyl) suberate
BSA	Bovine serum albumin
C	Cytosine (DNA) or Cysteine
C-	Carboxyl
cAMP	Cyclic adenosine monophosphate
cDNA	complementary DNA
CBB	Coomassie brilliant blue
cpm	Counts per minute
D	Aspartate
ddNTP	Dideoxynucleoside triphosphate
DEPC	Diethyl pyrocarbonate
DMS	3,3'-dimethyl suberimidate
DMSO	Dimethyl sulfoxide
DNA (agar)	Difco nutrient agar
DNA	Deoxyribonucleic acid
DNase	Dexoyribonuclease
ds	Double stranded
dNTP	Deoxynucleoside triphosphate
DTT	Dithiothreitol
E	Glutamate
ECF	Extracytoplasmic function
EDTA	Ethylenediamenetetraacetic acid
F	Phenylalanine
G	Guanine (DNA) or Glycine

GFP	Green fluorescent protein
GTP	Guanosine triphosphate
H	Histidine
HCV	Hepatitis C virus
His	Histidine
I	Isoleucine
IPTG	Isopropyl β-D-thiogalactopyranoside
K	Lysine
kb	kilobase
L	Leucine
LB	Luria Bertani medium
M	Methionine
MBP	Maltose-binding protein
MBSU	Molecular biology services unit
mRNA	Messenger RNA
MS	Mannitol Soy agar
MW	Molecular weight
MWM	Molecular weight marker
N	Asparagine
N-	Amino
nt	Nucleotide
NTG	Nitrosoguanidine
NTPase	Nucleoside triphosphatase
NTP	Nucleoside triphosphate
OD	Optical density
ORF	Open reading frame
P	Proline
PBP	Penicillin-binding protein
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
ppGpp	Guanosine tetraphosphate
pre-mRNA	precursor messenger RNA

Q	Glutamine
R	Arginine
R2YE	Sucrose yeast-extract medium
RBS	Ribosome binding site
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RNPase	Ribonucleoprotein displacement activity
rpm	Revolutions per minute
RT	Reverse transcriptase
RT-PCR	Reverse transcriptase polymerase chain reaction
S	Serine
SCB	*S. coelicolor* γ-butyrolactone
sco3550	*S. coelicolor* open reading frame 3550
SDS	Sodium dodecyl sulfate
SF	Superfamily
SSC	Standard saline citrate
ss	Single stranded
snRNA	Small nuclear RNA
snoRNA	Small nucleolar RNA
spp	species
T	Thymine (DNA) or Threonine
TAE	Tris-acetate EDTA buffer
TBS-T	Tris-buffered saline-tween buffer
TBE	Tris-borate EDTA buffer
TE	Tris-EDTA buffer
TEMED	Tetramethyl ethylene diamine
Tm	Melting temperature
tRNA	Transfer RNA
TSB	Trypticase soy broth
U	Unit (measure of enzyme activity)
UTR	Untranslated region
UV	Ultraviolet
V	Valine
W	Tryptophan
w/v	Weight per volume

x-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside
Y	Tyrosine

Chapter 1:

Introduction

1. Introduction

1.0 General Overview of the Streptomycetes

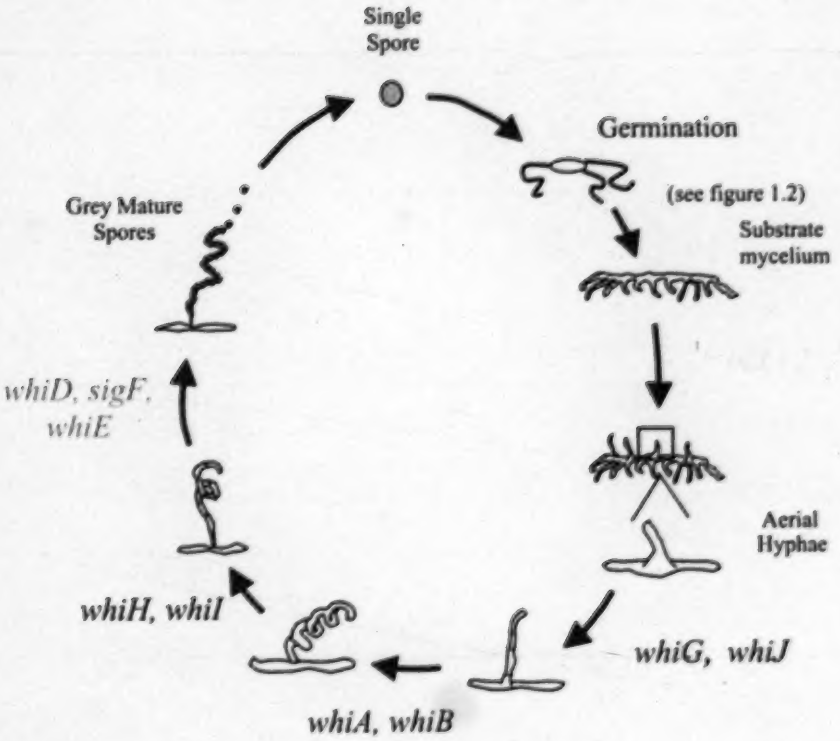
1.1 Physiological Adaptations for Survival in a Diverse Soil Ecosystem

Streptomyces species are obligately aerobic, filamentous, Gram positive eubacteria that are well recognized for their abundant distribution in soil and prolific generation of secondary metabolites, many of which have clinical relevance. Several key physiological adaptations have arisen in *Streptomyces* that allow the organism to survive and excel in the physical, nutritional and biological flux of diverse soil environments. Robust metabolic adaptations ensure initial survival of the streptomycetes and are exemplified by a lack of stringent growth requirements coupled with a high degree of metabolic flexibility. This is exemplified by the production and excretion of extracellular hydrolytic enzymes that aid in the biodegradation of complex plant and animal tissues. Metabolic resiliency in *Streptomyces* is complemented by growth as a branching vegetative mycelium, which adheres strongly to soil particle substrates in a hydrophilic environment, and creates a large surface area for soil nutrient utilization. Propagation in soil is supported by a developmental program, observed both in natural and laboratory settings, which aids in the effective production and dispersal of spore progeny. Although less resistant to environmental forces than endospores, *Streptomyces* spore progeny are hydrophobic and can survive over long periods of time due to resistance to desiccation, cold shock, anaerobic conditions and changes in osmotic pressure (Korn-Wendisch and Kutzner, 1992). Furthermore, the production of secondary metabolites, in particular antibiotics, from the developing mycelium may also impart a competitive advantage for streptomycetes over surrounding soil bacterial populations and further contribute to the prosperity of *Streptomyces spp.* in the soil (Champness, 2000b).

Streptomyces spp. are viewed as true multicellular organisms that undergo developmental transitions through distinct cell types (Miguelez et al., 2000; Bentley et al., 2002). The general morphological progression of a developing *Streptomyces* colony begins with the germination of a single, unigenomic spore. An emerging germ tube then grows, as a non-dividing hyphal tip, into a dense branching vegetative (substrate) mycelium which has little cross-wall septation and is multigenomic; therefore, a single hyphal compartment can contain up to 1000 nucleoids (Miguelez et al., 1999). Upon nutritional deprivation or stress related signals, a process of morphological differentiation manifests as the upward growth of thick-walled aerial hyphae that form as side branches upon the dense substrate hyphal network. A fibrous sheath surrounds the rarely branching aerial hyphae and presumably provides resistance to desiccation in order to support hyphal erection in a virtually non-aqueous environment. Morphogenesis of aerial hyphae also coincides with antibiotic production from the substrate hyphae, which potentially protects the developing organism from infringing microbial populations and may monopolize the limited nutrients in its sphere of growth. Finally, aerial hyphae undergo a coiling process and normal cell division to form septated compartments that culminate with the formation of hydrophobic, unigenomic spores that can now be dispersed to the surrounding environment (Figure 1.1) (Wildermuth, 1970; Plaskitt and Chater, 1995; Hopwood, 1999).

Developmental cell type transitions and differentiation to form aerial hyphae and spores are supported by several vital physiological traits. For instance, morphogenesis is partially fuelled by two distinct phases of glycogen deposition present in developing *Streptomyces* colonies. Glycogen stores are present in the substrate hyphae at the air-water interface at periods of aerial mycelium formation, while a second phase of

Figure 1.1 General Life Cycle of the Streptomycetes: *Streptomyces* spp. undergo an elaborate developmental program in order to produce spore progeny for dissemination and propagation in soil. Growth begins with the germination of a single spore which grows into a dense substrate mycelium. Developmental cues prompt morphogenesis and aerial hyphae formation. Aerial hyphae then begin sporulation through a process of coiling and septation which culminates in the formation of grey pigmented, hydrophobic spores. The *bld* genes (green) constitute a complex regulatory network that is believed to integrate environmental and intracellular cues to mediate growth phase transitions and initial hyphal erection. Subsequently, the early *whi* genes (blue) prime aerial hyphae for sporulation, and late *whi* genes (red) promote spore maturation and grey pigment production. (Modified from BK Leskiw original image)


Single
Spore
Germination
(see figure 1.2)
Substrate
mycelium
Aerial
Hyphae
whiG, whiJ
whiA, whiB
whiH, whiI
whiD, sigF,
whiE
Grey Mature
Spores

glycogen deposition is found in aerial hyphae tips during intermediary stages of sporulation. These distinct phases of glycogen accumulation appear to be under separate developmental regulation and are believed to facilitate morphogenesis by providing an accessible carbon and energy source. Also, the breakdown of glycogen stores provides smaller branched molecules that may be utilized by upwardly growing aerial hyphae as a means of regulating turgor pressure (Plaskitt and Chater, 1995).

The production of small, extracellular, hydrophobic molecules is also of paramount importance to *Streptomyces* morphogenesis. The hydrophobic, small, secreted peptide, SapB (spore associated protein B), forms a hydrophobic layer at the air-water interface in conjunction with a polymerized hydrophobic rodlet layer composed of RdlA and RdlB peptides, which eventually coat the aerial hyphae and spores (Willey et al., 1991; Claessen et al., 2002). The polymerized rodlet layer appears less important for morphogenesis and more involved in mediating hydrophobic surface attachment (Claessen et al., 2002). On the contrary, SapB dramatically lowers the surface tension of the air-water interface and fosters the upward growth of aerial hyphae from the hydrophilic environment of the substrate mycelium. Impetus for the erection of aerial hyphae appears to be provided by increasing turgor pressure and greatly facilitated by hydrophobic hyphal surfaces (Willey et al., 1991; Tillotson et al., 1998). Additionally, a family of small, hydrophobic, extracellular peptides has been identified in *Streptomyces* and designated chaplin (Chp) proteins (*coelicolor* hydrophobic aerial proteins). A subset of these chaplin proteins contain a sortase recognition motif which mediates peptide covalent attachment to the cell wall. Theoretically, cell wall associated chaplins may form a nucleation point for hydrophobic polymerization of non-cell wall associated small chaplin peptides and in combination with the hydrophobic SapB and rodlet layer, may

generate the classically described hydrophobic fibrous sheath surrounding emerging aerial hyphae and spores (Elliot et al., 2003a).

The balance between cell proliferation and cell death appears tightly regulated in *Streptomyces* development (Miguelez et al., 1999; Miguelez et al., 2000). Hyphal senescence was initially observed in non-sporulating aerial hyphae (Wildermuth, 1970). More recently, a microscopy study has indicated that *Streptomyces* cell death primarily consists of a highly regulated internal cell dismantling. Programmed hyphal death includes degradation of cytoplasmic components and nucleoids, fragmentation of the plasma membrane but, surprisingly, the maintenance of an intact cell wall (Miguelez et al., 1999). It was proposed that a region of hyphal death behind the emerging zone of aerial hyphae and sporulation would serve to conserve nutrients for the developing colony through the elimination of superfluous colony structures and also allow the recycling of nutrients and macromolecular components (Wildermuth, 1970; Miguelez et al., 1999). Furthermore, transitions in colony architecture, containing dead hyphae with intact cell walls, may act as both a source of mechanical support and as an elaborate conduit system to conduct the flow of recycled compounds to growing regions of the colony (Miguelez et al., 1999; Miguelez et al., 2000). Clearly, the physiological components, effectors, and overall role of the *Streptomyces* developmental program has responded and adapted to the diverse soil ecosystem in which the organism inhabits.

Secondary metabolite production, such as antibiotics, pigments, signaling molecules and siderophores play a variety of roles to enhance *Streptomyces* growth and survival (Nguyen et al., 2002). For example, the grey pigment inherent in mature spores is a polyketide secondary metabolite that confers UV resistance and the protection of progeny DNA from undue damage (Davis and Chater, 1990; Kelemen et al., 1998).

Also, geosmin, a terpenoid metabolite, which is in part responsible for the characteristic odour of soil, may act as a developmental signaling molecule (Gust et al., 2003). Of particular commercial interest, physiological differentiation of the substrate hyphae prompts the production and excretion of antibiotics that presumably provide a zone of inhibition to adjacent bacterial populations and present a competitive advantage to the developing colony. *Streptomyces* species represent a vast source of greater than two-thirds of the world's naturally derived antibiotics including pharmaceutically relevant compounds, such as anti-microbial, anti-tumor, anti-parasitic and anti-fungal agents, herbicides, insecticides and immunosuppressants. Consequently, there is thriving interest in the elucidation of antibiotic functions and biosynthetic pathways that may ultimately lead to industrial-level, novel and hybrid antibiotic production (Champness, 2000a; Watve et al., 2001; Thompson et al., 2002).

1.2 Completion and Extensions of the *Streptomyces coelicolor* Genome Sequence

The completion and initial annotation of the *Streptomyces coelicolor* A3(2) genome sequence has recently provided a greater insight into the genomic organization and content necessary to generate the manifold physiological adaptations that define *S. coelicolor* growth and survival in soil ecosystems. The *S. coelicolor* genome is linear, GC rich, and is large in comparison to other bacterial genomes, consisting of 8.7 megabases and predicted to contain 7,825 genes. The genomic organization of the linear chromosome reveals a central core that is nearly half the size of the genome and encodes proteins that are essential to normal growth processes: cell division; DNA replication; gene and protein expression; and amino acid biosynthesis. In contrast, right and left chromosomal arms flank the central core and contain genes potentially responsible for

physiological flexibility, such as secondary metabolite gene clusters and genes encoding hydrolytic extracellular enzymes. It is believed that these arms have expanded through lateral gene acquisition and intrachromosomal gene transfer and duplication, most likely taking different evolutionary routes among separate *S. coelicolor* species. It was noted that gene duplications may represent cell type specific isoforms, a novel feature in prokaryotes, where paralogous genes would be expressed at different stages in development (Bentley et al., 2002). This hypothesis is supported by the presence of duplicate anabolic operons responsible for the separate, distinct phases of glycogen deposition in the substrate and aerial hyphae (Bruton et al., 1995).

A general assessment of the annotated genes revealed a high degree of potential regulatory proteins representative of greater than one-tenth of the genome. These include a staggering 65 sigma (σ) factors, which indirectly confirms the important role of differential gene expression in the *S. coelicolor* life cycle. Notably, annotation predicted that 45 of the 65 sigma factors are the aptly named extracytoplasmic function σ factors that respond to extracellular stimuli. Annotation of the genome sequence also identified 85 sensor kinases and 79 response regulators, with 53 of these serving as sensor-regulator pairs involved in two-component signal transduction. This is coupled with the prediction of 614 genes encoding proteins with a transport function and 819 genes encoding putative secreted proteins, thereby emphasizing the importance of sensing and responding to external stimuli in a highly variable environment (Bentley et al., 2002). The identification of 20 novel secondary metabolite clusters, which may function in a physical, chemical or biological stress responsive role, were predominately localized to a region by the right-hand core-arm boundary. This is combined with the discovery of 30 gene clusters predicted to be involved in secondary metabolite synthesis from the recent

annotation of the *Streptomyces avermitilis* genome, a species of pharmaceutical interest due to its production of avermectins that act as anti-parasitic agents, and reaffirms the legitimacy of continued commercial interest in the streptomycetes as a source of novel secondary metabolites (Bentley et al., 2002; Nguyen et al., 2002; Ikeda et al., 2003).

The completed genome sequence now allows researchers to circumvent classical single gene hypothesis-driven research with whole genome analysis, in order to enhance the current understanding of regulatory networks involved in development and antibiotic production. Extensions of the *S. coelicolor* genome sequence and a movement into the post-genomics era of high-throughput genome-wide assays are evident in several preliminary *Streptomyces* studies (Vohradsky et al., 2000; Huang et al., 2001; Hesketh et al., 2002b). For example, two-dimensional gel electrophoresis allows separation and resolution of the majority of proteins produced by a cell at a given time point (Hesketh et al., 2002a). Proteomic comparisons in *Streptomyces*, assisted by mass spectrometry peptide identification technology, have revealed an apparent link between protein stress stimulons, heat shock and osmotic shock in particular, with development (Vohradsky et al., 2000). Also, two-dimensional gel electrophoresis has shown prevalent regulation of peptides through extensive post-translational modification in both vegetative growth and developing mycelia and has led to a predicted value of 1.2 proteins per gene (Hesketh et al., 2002a). Microarray analysis, a hybridization *in silico* technique for examining global gene expression changes, has contributed to the understanding of growth phase responsive gene expression and the identification of clusters of coordinately regulated genes involved in antibiotic production (Huang et al., 2001). Consequently, the *Streptomyces* post-genomic era has already heralded a profusion of data concerning gene and protein regulatory interactions within a complex developmental network.

1.3 Genetic Characterization of Morphological Differentiation

Growth phase transitions and the developmental complexity of *Streptomyces* have prompted a great deal of focus on the genetic characterization of morphological and physiological differentiation. *Streptomyces coelicolor*, the most genetically well characterized streptomycete, was originally used in genetic studies due to its production of two pigmented antibiotics for use as genetic markers (diffusible, blue-pigmented actinorhodin and cell-associated, red-pigmented undecylprodigiosin) (Hopwood, 1999). Two primary classes of developmental mutants have been isolated in *S. coelicolor*, with some homologs in other streptomycetes. First, mutants that are blocked at the earliest stage of sporulation and unable to produce or erect aerial hyphae are denoted as having a *bld* (bald) phenotype and visualized as smooth colonies lacking white, fuzzy aerial hyphae (Merrick, 1976). Second, mutants that lack complete morphogenesis of aerial hyphae into spores are designated as *whi* (white) mutants due to the absence of the characteristic grey pigmentation of mature spores (Chater, 1972). Furthermore, genetic characterization has revealed that aerial hyphae production and sporulation are controlled through separate regulatory networks that are defined by several emerging features of prokaryotic development, such as temporal and spatial differential gene expression, developmental checkpoints, and three-dimensional protein localization (Figure 1.1) (Kroos and Maddock, 2003).

1.4 *whi* Mutants

whi mutants are impaired in sporulation and were originally isolated as colonies that had wild-type aerial hyphae production and yet were completely blocked in the formation of grey pigmented mature spores (Chater, 1972). The normal sequence of sporulation consists of the growth of aerial hyphae by cell wall extension at hyphal tips

and a subsequent halt in growth, which coincides with initial sporulation, septation and coiling. This is followed by spore maturation which includes nucleoid condensation and partitioning into spore compartments, hydrolysis of aerial hyphae glycogen stores, formation of a series of spore-specific cell wall layers, a general rounding of septated spores, grey pigment production and eventual fragmentation of mature spore chains (Wildermuth and Hopwood, 1970; Chater, 1998; Flardh et al., 1999). A study of single and double *whi* mutant phenotypes noted a variety of *whi* mutant classes impaired at sequential stages of sporulation. These phenotypes ranged from mutants containing aerial hyphae that were completely undifferentiated with straight hyphal stalks, those with partially coiled hyphae or irregular septation, and strains with differentiated hyphae and elongated, aberrant spore structures or nearly wild-type ellipsoidal spores which lack only proper pigmentation (Chater, 1972; Chater, 1975). Characterization of mutants revealed that the *whi* genes all encode putative regulatory products and the inability to uncover key structural or enzymatic effector components in a genetic screen may imply some inherent degree of functional redundancy in these systems (Chater, 1998). Based upon *whi* single and double mutant phenotypes, the *whi* genes were placed in an epistatic pathway where early *whi* genes that are integral to the primary septation decision and cessation of aerial hyphae growth are essential for the induction of the late *whi* genes, which control the final maturation of spores (Chater, 1975; Ryding et al., 1999).

The early *whi* genes are expressed shortly after aerial hyphae formation and consist of two separate regulatory pathways that are critical for a process of sporulation-specific septation driven by normal cell division within aerial hyphae (Chater, 2001). *whiG* mutants produce straight, undifferentiated aerial hyphae (Chater, 1972; Chater et al., 1989). Moreover, *whiG* is expressed independently of other early *whi* genes and is

thereby placed at the beginning of the *whi* gene hierarchy (Kelemen and Buttner, 1998). Consistent with developmental models that predict a role for differential gene expression in stages of development, *whiG* encodes a sporulation- specific σ factor (Chater et al., 1989). Homologs of σ^{whiG} exist in nearly all *Streptomyces* species and, interestingly, show high similarity to a motility σ factor (σ^{28} or σ^{D}), rather than a sporulation σ factor, of *Bacillus subtilis* (Tan and Chater, 1993). The corresponding *whiG* gene is expressed constitutively throughout growth, which suggests the possibility of post-translational regulation, perhaps through binding of an anti-σ factor. Release of the predicted anti-σ factor and liberation of σ^{whiG} would most likely occur at the onset of aerial hyphae formation. Indicative of the importance of *whiG* gene dosage to promote entry into sporulation, an overproduction of *whiG* in vegetative hyphae prompted a degree of ectopic septation, while the addition of competing σ^{28} promoter sequence caused a decrease in sporulation, potentially due to σ^{whiG} sequestration (Chater et al., 1989).

The *whiG* encoded σ-factor is directly responsible for the expression of a ProX protein, which is not essential for sporulation, but is possibly involved in glycine betaine binding and the regulation of hyphal turgor pressure (Tan et al., 1998). σ^{whiG} also directs expression from at least two early *whi* genes, *whiH* and *whiI*, which encode putative regulatory factors (Ryding et al., 1998; Ainsa et al., 1999; Flardh et al., 1999). *whiH* and *whiI* have similar morphological mutant phenotypes consisting of loosely coiled, wavy aerial hyphae (Chater, 1972; Ryding et al., 1998; Ainsa et al., 1999). Initial characterization of *whiH* and *whiI* single and double mutants revealed that *whiI* mutants had significantly higher fragmentation than *whiH* mutants, which is indicative of more spore septation and therefore *whiH* was designated epistatic, or functional earlier in the *whi* gene network, to *whiI* (Chater, 1975; Ryding et al., 1999). Products of both genes

possibly sense nutritional or external stimuli as *whiH* encodes a potential repressor protein with similarities to a large family of metabolic response repressor proteins, and *whiI* encodes a predicted atypical response regulator. Unlike other response regulator genes, *whiI* has no known genomic link to a putative sensor kinase and encodes a response regulator containing an unconventional phosphorylation pocket. Both genes appear to be σ^{whiG} dependent with maximum levels of expression during sporulation (Ryding et al., 1998; Ainsa et al., 1999). However, the finding of *whiH* epistasis is inconsistent with more recent observations of nucleoid partitioning in the early *whi* mutants. At early stages of sporulation, up to fifty uncondensed nucleoids can be distributed along the length of wild-type aerial hyphae, which have interspersed vegetative crosswalls. Sporulation-specific septation is the only true cell division that occurs in the *Streptomyces* life cycle because vegetative and aerial hyphae grow through tip elongation and branching and have rare cross-walls. Cell division comprises condensation of nucleoids, in conjunction with partitioning and septum formation in order to form unigenomic spores. The state of nucleoid partitioning was examined in early *whi* mutants and revealed that *whiG* (and *whiA* and *whiB*) mutants had no visible signs of nucleoid dynamics or septation while *whiH* mutants showed some sporulation septa with irregular regions of delineated, condensed DNA. Interestingly, *whiI* mutants displayed a similar physiological phenotype to *whiG* (and *whiA* and *whiB*) mutants, consisting of non-partitioned, uncondensed nucleoids, which is indicative of *whiI* epistasis (Flardh et al., 1999). To complicate matters further, it has been observed early in sporulation that both *whiH* and *whiI* negatively autoregulate themselves, and there is also evidence for *whiI* negative cross-regulation of *whiH*. Also, disruption of *whiH* yields strains with minor grey pigment and diminished *whi* gene expression, while *whiI*

disruption completely abolishes late *whi* gene expression. A speculative model predicts that *whiH* and *whiI* respond to sporulation specific signals, which induce modification and relieve auto-regulation and cross-regulation; this leads to a burst of WhiH and WhiI production, which in turn activates genes responsible for spore septation, late *whi* gene activation and ensuing spore maturation (Ryding et al., 1998; Ainsa et al., 1999; Flardh et al., 1999).

Independent of σ^{whiG} gene activation, *whiB* and *whiA* comprise a separate early *whi* gene pathway which, in addition to the σ^{whiG} dependent pathway, serves as an essential checkpoint before entry into late phases of sporulation. *whiB* encodes an 87 amino-acid cysteine-rich cytoplasmic protein, which possibly senses redox changes WhiB shows some sequence similarity to transcription factors and has several expressed paralogs, designated *whiB*-like (*wbl*) genes, in *S. coelicolor* (Davis and Chater, 1992; Soliveri et al., 2000). *whiA* encodes a peptide of unknown function, although homologues have been found in all gram-positive prokaryotes. *whiA* is in the same early *whi* gene hierarchial group as *whiB*; thus presenting a mutant phenotype with abnormally long, coiled aerial hyphae bereft of regular septation, nucleoid condensation or partitioning (Chater, 1972; Ainsa et al., 2000). Early epistatic studies of mutant morphology, which are now regarded as not completely representative of the *whi* gene network, place *whiA* and *whiB* downstream of *whiG* and upstream of *whiH* and *whiI* (Chater, 1975; Flardh et al., 1999). Yet, epistatic studies lack the ability to define parallel pathways. Whereas the σ^{whiG} dependent pathway is believed to prime the aerial hyphae for sporulation through activation of coiling and cell division, *whiA* and *whiB* are believed to respond to sporulation-specific signals and prompt a cessation of aerial hyphal growth. Potentially, a stoppage in growth allows the ordered septation process

mediated by *whiH* (and potentially *whiI*) to form regular septated spore compartments (Flardh et al., 1999). To this point, there has been no observed direct interaction between the *whiG* and *whiA/whiB* sporulation pathways, however, it is possible that *whiA* and *whiB* may be responsible for *whiH* and *whiI* modification, or that the cessation of aerial hyphae growth may modulate proper positioning or accumulation of other early *whi* gene products.

Upon cessation of aerial hyphal growth and sporulation septa formation, activation of the late *whi* genes promotes spore maturation and grey pigment production occurs. Spore maturation consists of a progressive thickening of the spore cell wall, which entails the synthesis of new cell wall components and modification of the existing cell wall. Furthermore, once septa are present and form cylindrical spore compartments that contain loosely condensed nucleoids, the initial layer of the spore wall is constructed inside the remaining hyphal cell wall. Generation of a second spore wall layer molds the compartment into a more ovoid shape and coincides with complete chromosomal condensation, pigment production and spore fragmentation (Wildermuth and Hopwood, 1970; Potuckova et al., 1995).

Late stages of spore differentiation are dependent upon gene expression directed by a separate σ factor, designated σ^F, which was originally identified in a search for σ factors in *Streptomyces aureofaciens* and later confirmed to have homologs in *S. coelicolor* and other streptomycetes. σ^F is a member of the well-conserved σ^{70} family and shows some identity to the –35 and –10 promoter recognition motifs of σ^B, a *B. subtilis* σ factor involved in the general stress response and stationary phase gene expression (Potuckova et al., 1995). The position of *sigF* in the *whi* gene network occurs at later stages of sporulation, and *sigF* expression is absolutely dependent upon *whiA*,

whiB, *whiG* and *whiI*, and to a lesser degree, *whiH* and *whiJ*. As well, *sigF* expression appears coincidentally with septum formation, which most likely ensures a proper sequence of sporulation events (Kelemen et al., 1996). *sigF* mutants form small, aberrant, thin walled spores that are unusually sensitive to detergents and less subject to fragmentation than wild-type spores. Chromosomes also lack complete condensation, which is possibly an artifact of a thinner cell wall and the resultant permeability increase to fixation compounds used in microscopy (Potuckova et al., 1995). Although a simple model of sporulation would place σ^F expression under the control of σ^{whiG}, this does not appear to be the case, nor does the *whiH* encoded transcription factor appear to bind near the *sigF* promoter region. Consequently, *sigF* expression appears to be indirectly regulated by the products of *whiG* and *whiH*, potentially through an as yet unidentified σ factor, with other early *whi* genes acting as candidate transcription factors for either direct or indirect regulation (Kelemen et al., 1996; Homerova et al., 2000).

Targets of *sigF* are likely to play a role in spore wall synthesis such as spore specific penicillin-binding proteins (PBPs) and other spore-associated proteins. One possible *in vivo* target of the σ^F RNA polymerase holoenzyme is the *whiEP2* gene, which is part of the *whiE* locus. This is a terminal *whi* gene cluster responsible for the type II polyketide synthesis, similar to aromatic polyketide antibiotic synthesis, of the grey spore pigment. The *whiE* locus consists of two divergent promoters: *whiEP1*, which directs expression of a seven gene operon; and *whiEP2*, expressing a single gene. Expression from both promoters is dependent upon all of the early *whi* genes, and *whiEP2* activation is dependent upon σ^F activation. However, *in vitro* transcription did not confirm σ^F holoenzyme binding, and thus the *in vivo* effect may be indirect or due to additional interactions of another transcription factor (Kelemen et al., 1998).

A further identified late *whi* gene is *whiD*, which encodes a product showing homology to the *whiB* family of transcription factors and may act as a redox stress sensor through intramolecular or intermolecular disulphide bond formation (Soliveri et al., 2000). The regulation of *whiD* is still unclear; however, expression coincides with the appearance of sporulation septa. The *whiD* product is postulated to play a role in spore maturation and may also affect the initiation of sporulation septation. This function is evident by the *whiD* mutant phenotype that has reduced levels of sporulation and generates spores that are consistently heat sensitive and highly irregular in size. Also, the mutant strain has irregular septum formation and spore wall deposition, which contributes to the presence of small compartments that completely lack chromosomal DNA (Molle et al., 2000).

Clearly, a complex regulatory network, directed by the *whi* genes, is responsible for the morphogenesis of aerial hyphae into mature spores that are capable of environmental dissemination as resilient progeny of the streptomycetes. While the initial framework of the regulatory circuitry has been determined, a lack of knowledge still exists for the environmental and developmental signals that are sensed by the *whi* gene products, as well as the enzymatic and structural targets of the *whi* gene-encoded regulatory factors (Chater, 1998; Chater, 2001). The advent of *S. coelicolor* post-genomic analysis will surely provide a valuable resource for the identification of subordinate targets while the use of single cell reporter constructs, such as enhanced green fluorescent protein fusions, may provide finer knowledge of the localization of gene expression and protein products in the developing aerial mycelium.

1.5 *bld* Mutants- a Link to Stress, Metabolism and Development

The *bld* (bald) mutants comprise a second class of *Streptomyces* developmental mutants, which exhibit a pleiotropic block in aerial mycelium formation and antibiotic production. The temporal regulation and excretion of SapB, an abundant, small, hydrophobic morphogen located in spore walls and in a diffusible zone around developing colonies, is essential to aerial hyphae production (Willey et al., 1991). Interestingly, *bld* mutants evidently lack SapB production, and the application of purified SapB to *bld* mutants or the juxtaposition of *bld* mutants with SapB producing wild-type strains results in extracellular complementation and a restoration of aerial hyphae formation, but not antibiotic production. Although SapB is essential for morphogenesis and SapB production is absent from *bld* mutants, the exogenous addition of SapB exacts only transient aerial hyphae production with no ensuing sporulation, and suggests other developmental defects are present in the aerial hyphae of *bld* mutants (Willey et al., 1993). The action of SapB as a morphogenic effector of aerial hyphae formation with no other regulatory role is also supported by successful attempts at restoring transient aerial hyphae production to *bld* mutants through the exogenous addition of a *Schizophillum* hydrophobic protein involved in fungal morphogenesis (Tillotson et al., 1998). Recently, there is evidence that *ramS*, a gene included in the *ram* gene cluster, which is a complex locus encoding five gene products that cooperatively facilitate aerial hyphae formation and effectively uncouple morphological differentiation from antibiotic production, encodes the SapB morphogen (Keijser et al., 2002; Nguyen et al., 2002) (Kodani et al., 2004).

The *ram* (rapid aerial mycelium) gene cluster was originally identified as a *S. coelicolor* clone that, when overexpressed on a low copy number vector, accelerated aerial mycelium formation in developmentally delayed strains of *Streptomyces lividans*.

Clones that stimulated development were sequenced and the corresponding genes were identified as *ramA* and *ramB*, encoding putative ATP-dependent membrane translocases, and *ramR*, situated downstream of *ramAB* and encoding a putative UhpA-like response regulator (Ma and Kendall, 1994). Two additional *ram* genes have also been identified to date: *ramC*, encoding a putative serine/threonine sensor kinase, and *ramS*, encoding a putative small signaling molecule (Keijser et al., 2000; O'Connor et al., 2002). Moreover, *ramCSAB* exists in an apparent operon while downstream *ramR* is expressed as a monocistronic transcript (Keijser et al., 2000). A homolog of the *S. coelicolor ram* gene cluster is highly conserved and expressed in *S. lividans*, whereas the *ram* gene cluster shows protein homology, with little DNA sequence identity, to the *amf* gene cluster which is important for aerial hyphae production in *Streptomyces griseus* (Ma and Kendall, 1994). Overexpression of *ramAB* and *ramR* shows negligible effects in *S. coelicolor*, and the mode of stimulation of aerial hyphae production in *S. lividans* is not easily explained. A *ramAB* null mutant also completely abolishes hyphal morphogenesis in *S. lividans* while *S. coelicolor ramAB* mutants retain their normal development. Consequently, the production of hyphal morphogenesis in *S. lividans* appears to rely completely upon the *ram* gene cluster, while *S. coelicolor* potentially contains an auxiliary *ram*-independent pathway (Keijser et al., 2000).

Further analysis of the *ram* gene cluster has revealed an upregulation of *ramR* gene expression corresponding to initial aerial hyphae formation and a direct dependence of *ramCSAB* expression upon *ramR* expression, presumably through upstream promoter binding of RamR. *ramR* and *ramCSAB* expression are also completely abolished in several *bld* mutants that fail to produce SapB, whereas *ramR* and *ramCSAB* expression is

wild-type in the earliest *whi* mutants (Keijser et al., 2002). The hypothetical link between the *ram* gene cluster and SapB production is further substantiated by the finding that overexpression of *ramR* in *bld* mutant backgrounds restores aerial hyphae production through the apparent upregulation of SapB production. In addition, *ramR* mediated SapB production also requires the expression of *ramC* and *ramS* and thus *ramR*, *ramC* and *ramS* are effectively a class of *bld* genes since all appear essential for morphological differentiation of *S. coelicolor* on rich medium (Nguyen et al., 2002; O'Connor et al., 2002).

Until recently there was little genetic characterization of the SapB gene since the purified 18 amino acid peptide was presumably glycosylated and the adjoining sugar moiety interfered with attempts at Edman degradation amino acid sequencing (Willey et al., 1991). Yet, comparison of small portions of the deduced *ramC* encoded protein, which was previously believed to be a serine/threonine kinase, to the databases revealed a (C-terminal) region similar to enzymes involved in the biosynthesis of lantibiotics, which are oligopeptide antibiotics that are ribosomally synthesized in Gram positives as inactive precursors that are post-translationally modified by the formation of lanthionine bridges. Post-translational modification of lantibiotic prepeptides customarily involves proteolysis and the dehydration of serine and threonine residues to form didehydroalanine (Dha) and didehydrobutyrine (Dhb) that may react with sulfhydryl groups of cysteine residues to form acid-stable thioether lanthionine bridges (or in the case of Dhb, 3-methylanthionine bridges). Free Dha or Dhb that have not reacted with sulfhydryl groups can block Edman degradation and it was hypothesized that RamC may process RamS as a lantibiotic-like peptide and possibly lead to the mature SapB product. Furthermore, the inability to

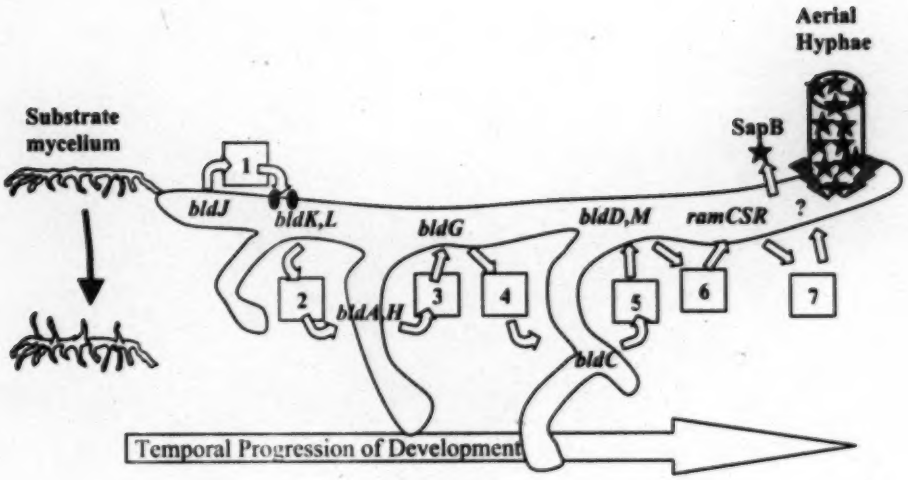
identify the amino acid sequence of SapB through Edman degradation could have been due to Dha blockage at a dehydrated serine residue. These suppositions were confirmed by structural analysis of SapB where chemical modification was employed to modify Dha and Dhb residues in an effort to make the peptide more susceptible to Edman degradation. The results revealed that the amino acid sequence of SapB was identical to the predicted C-terminal sequence of RamS (Kodani et al., 2004).

Curently, a speculative model for SapB-dependent morphogenesis proposes that the *bld* gene regulatory network serves to integrate stress response and nutritional signals within a signaling pathway, with the endpoint being the induction of RamR as a key effector of hyphal morphogenesis and the direct activator of *ramCSAB* expression. Expression of RamC may potentially lead to processing of RamS to form the mature lantibiotic-like peptide SapB morphogen, which may achieve plasma membrane transport through the RamAB exporter protein and possibly other transporters with overlapping function. Finally, the *ram* gene cluster may activate or function with parallel pathways that ensure complete morphogenesis and the *ram* gene cluster may in fact reside as a final developmental checkpoint for morphological differentiation and the formation of aerial hyphae (Chater and Horinouchi, 2003). Finally, a possible SapB-independent pathway for aerial mycelium formation is suggested by the observation of small amounts of extremely delayed aerial hyphae production in a *ramR* mutant, and the fact that *bld* mutants grown on minimal mannitol medium show restored aerial hyphae formation without *ram* gene expression or SapB production (Merrick, 1976; Champness, 1988; Nguyen et al., 2002).

The restoration of wild-type morphological differentiation to *bld* mutants, when plated adjacent to wild-type *S. coelicolor* strains, was further elaborated upon by the discovery of extracellular complementation between certain pairs of *bld* mutants. It was found that the juxtaposition of certain *bld* mutants lead to the restoration of SapB production, recovery of aerial hyphae production and sporulation to one of the *bld* mutants, although to a weaker and more delayed extent that that of wild-type complementation. Extracellular complementation between *bld* mutants was postulated to be a manifestation of an extracellular signaling cascade where separate *bld* mutants were responsible for the production of different signals in a pathway, which would ultimately converge upon effectors of SapB production (Willey et al., 1993). A constraint of this hypothesis requires *bld* mutants to unidirectionally complement upstream *bld* mutants through the diffusion of signaling molecules which allows the upstream mutant to bypass the pathway step in which they are blocked. Conversely, upstream mutants cannot complement downstream mutants as they are unable to produce the desired signal that is deficient in the downstream *bld* mutant. Extracellular complementation studies have described eight *bld* gene complementation groups requiring a minimum of seven extracellular signals in a proposed *bld* gene intercellular signaling cascade: [*bldJ*]→[*bldK,bldL*]→[*bldA,H*]→[*bldG*]→[*bldC*]→[*bldD,bldM*]→[*ramCSR*]⇒SapB (Willey et al., 1993; Nodwell et al., 1996; Nodwell et al., 1999; Nguyen et al., 2002). For instance, *bldJ* would be blocked at the earliest step of the signaling pathway and complemented extracellularly by all other *bld* mutants. A *bldK* mutation would complement a *bldJ* mutant only and would be complemented by all downstream *bld* mutants, and so on. As low concentrations of signal most likely diffuse throughout the medium, complementation only occurs at the fringes of the juxtaposed growth (Willey et

al., 1993). Even once effectively complemented, upstream *bld* mutants, with aerial hyphae production restored at the periphery of colony growth, will not likely produce enough diffusible signal to rescue a juxtaposed downstream mutant (Figure 1.2). In a signaling cascade of this nature, the *bld* genes may integrate developmental cues such as population density or nutrient depletion, and utilize extracellular signals to delineate various physiological checkpoints involved in assessing the status of vegetative hyphae. Each signaling complementation group may also represent a successive stage in developmental readiness and may correspond to a series of physiological processes leading up to complete hyphal differentiation (Figure 1.2) (Chater and Horinouchi, 2003). Observations that *bldB*, *bldI*, and *bldN* do not fit into the proposed signaling hierarchy and that antibiotic production is not recovered by extracellular complementation indicates that the regulatory scheme governing the *Streptomyces* developmental decision is probably more complex than that of a simple signaling pathway (Willey et al., 1993; Pope et al., 1998; Nodwell et al., 1999; Bibb et al., 2000). Alternative explanations for the phenomenon of extracellular complementation abound. It has been reasoned that these extracellular signals may be regulators or precursors in a biosynthetic pathway for the end-point production of SapB or a SapB activation signal (Nodwell et al., 1999). Another hypothesis predicts that *bld* genes may mediate various transitions in secondary metabolism and rather than involving a signaling mechanism, extracellular complementation may be a result of downstream mutants invoking chemical changes in the medium that bypass the developmental defect of upstream *bld* mutants. This hypothesis is based upon the apparent link between development and metabolism and the observation of carbon catabolite repression defects in most *bld* mutants (Pope et al., 1996; Karandikar et al., 1997). Moreover, a citrate synthase gene, *citA*, is essential for

Figure 1.2 The Proposed *bld* Gene Extracellular Signaling Cascade. Observations of extracellular complementation precipitated the belief that the *bld* genes act in a signaling pathway that ultimately converges upon the production of SapB morphogenic peptides (stars). The simplest incarnation of the proposed signaling cascade is the import of the *bldJ* encoded signaling molecule by the BldK transporter protein (ovals). The majority of *bld* genes encode regulatory proteins and presumably, indirectly promote the production of extracellular signaling molecules. There are seven putative extracellular signals that function in the cascade, where members of each *bld* gene complementation group are responsible for sensing the previous signal and generating the next signal in the pathway. The endpoint of the signaling pathway yields the production of hydrophobic peptides, such as SapB, the rodlins (lightning bolts) and the chaplins, which form a fibrous sheath (shaded cylinder) around emerging aerial hyphae to facilitate hyphal morphogenesis. Adapted from (Chater, 1998).


Aerial
Hyphae
Substrate
mycelium
SapB
1
bldJ
bldK,L
bldG
bldD,M
ramCSR
?
2
bldA,H
3
4
5
6
7
bldC
Temporal Progression of Development

proper functioning of the citric acid cycle and integral to proper development. *citA* mutants appear *bld* while, similar to the situation with other *bld* mutants, the growth medium becomes increasingly acidified (Viollier et al., 2001a). cAMP, and the *relA*/ppGpp-mediated stringent response to amino acid starvation is also important for secondary metabolic transitions and developmental processes in *S. coelicolor* (Chakraburtty and Bibb, 1997; Susstrunk et al., 1998). Possibly, global *bld* gene-mediated carbon catabolite repression and the stringent response allow transduction of metabolic stress and nutrient changes, which represents an important physiological checkpoint integrated within the *S. coelicolor* developmental circuitry (Pope et al., 1996; Karandikar et al., 1997; Viollier et al., 2001a; Viollier et al., 2001b).

Many *bld* genes have little characterization except for a description of their mutant phenotype and position in the proposed extracellular signaling cascade, such as *bldL*, *I* and *C* (Harasym et al., 1990; Nodwell et al., 1999). There is also little information pertaining to *bldB* but it is defined as the most pleiotropic *bld* gene. Like other *bld* mutants, *bldB* causes defects in morphological and physiological differentiation, extracellular signaling and carbon catabolite repression; however, in contrast to the majority of other *bld* mutants, hyphal morphogenesis is not restored to *bldB* mutants when grown on minimal medium containing mannitol as an alternate carbon source (Harasym et al., 1990; Pope et al., 1998). Preliminary characterization of *bldB* revealed homology to DNA binding proteins and exhibited an apparent developmental expression program, with maximal levels of transcription at late vegetative time points (Pope et al., 1998). BldB is postulated to bind to DNA as a dimer and negatively autoregulate its own expression through an indirect mechanism (Pope et al., 1998; Eccleston et al., 2002). A recovery of aerial hyphal growth by all other *bld* mutants when grown on minimal

medium with mannitol suggests that growth on mannitol stimulates a separate developmental network independent of *bld* gene regulation. Another possibility is that growth on mannitol can bypass the necessity for developmental checkpoints involving metabolic perception, such as carbon catabolite regulation and a response to medium acidification. More to the point, the inability of *bld* mutants to repress alternate catabolic pathways may be linked to an inability to trigger aerial hyphae formation under nutrient stressed conditions, and this is somehow circumvented through growth on mannitol. The extreme pleiotropy of the *bldB* gene may implicate BldB as a global regulator for both *bld* gene dependent development and *bld* gene independent development on mannitol, or for a separate developmental checkpoint for growth on a nonacidogenic carbon source (Champness, 1988; Pope et al., 1996; Pope et al., 1998).

The simplest explanation for extracellular complementation is that all *bld* genes encode either signaling molecules or transporter proteins in a *bld* gene signaling cascade. The direct association of a developmental signaling mechanism governed by some *bld* genes has been substantiated by studies of *bldJ* (*bld261*) and *bldK*. As shown above, these genes occur early in the proposed *bld* gene signaling pathway. Studies of *bldJ* and *bldK* mutants grown on media pre-conditioned by *bld* mutants have tentatively assigned *bldJ* as encoding a small, signaling oligopeptide (Nodwell et al., 1996; Nodwell and Losick, 1998). Similar to SapB, the BldJ putative oligopeptide is potentially covalently modified and is resistant to Edman degradation amino acid sequencing, while the *bldJ* gene has yet to be cloned (Nodwell and Losick, 1998). The import of the *bldJ* encoded signal has been exhibited to depend upon expression and assembly of the *bldK*-encoded subunits of a putative transmembrane ATP- binding cassette (ABC) transporter protein (Nodwell et al., 1996; Nodwell and Losick, 1998). The response to BldJ oligopeptide

transport, as the putative first signal in the cascade, and production of a putative second extracellular signal would then be dependent upon *bldA* or *bldH* within the third complementation group (Nodwell and Losick, 1998). Possibly, *bldJ* may constitute an initial quorum sensing- like signal which is transported through the BldK ABC transporter as an indicator of cell density and primes the vegetative hyphae to respond to other environmental cues. Yet, it is apparent that the majority of characterized *bld* genes do not encode signaling molecules or transporter proteins, but rather encode regulatory proteins with targets implicated in pleiotropic roles in *Streptomyces* physiology and development and thus, may affect extracellular signaling indirectly (Chater, 2001).

Besides signaling peptides, classes of hormone-like γ-butyrolactones may function in the *bld* gene signaling pathway. The supporting evidence comes from *S. griseus*, a distant relative of *S. coelicolor*, where A-factor, a γ-butyrolactone, is essential for morphogenesis and mediates its effects through ArpA, a cytoplasmic receptor protein (Khokhlov et al., 1967; Lezhava et al., 1997). In contrast, *S. coelicolor* lacks a close A-factor homolog and a similar class of *S. coelicolor* γ-butyrolactones (SCB) does not appear to be directly involved in development (Takano et al., 2000; Takano et al., 2001). However, in *S. coelicolor*, CprA, an orthologue of *S.griseus* ArpA, is involved in development and is indirectly essential for expression of the *ram* gene cluster (Onaka et al., 1998; O'Connor et al., 2002). Furthermore, *cprA* mutants are developmentally delayed and therefore, the impact of γ-butyrolactone signaling upon *S. coelicolor* development is expected to be significant (Onaka et al., 1998). Differences in γ-butyrolactone signaling between the highly diverged *S. coelicolor* and *S. griseus*, as well as other developmental homologs with separate developmental regulatory connections, is

hypothesized to be a result of adaptation to diverse soil ecosystems and a driving force for ongoing speciation in the streptomycetes (Chater and Horinouchi, 2003).

bldA is one of the most well characterized *bld* genes and is described as encoding a leucyl-tRNA, which is required for the efficient translation of transcripts containing the rare UUA codon. Furthermore, *bldA* mutants are devoid of both aerial hyphae and antibiotic production and, due to the paucity of TTA codons in GC rich *Streptomyces* DNA, most likely represent a level of translational developmental regulation (Lawlor et al., 1987). *bldA* is also the only identified *S. coelicolor bld* gene homolog in *S. griseus*, and this conservation, coupled with its developmental role, implicates *bldA* as encoding an integral global regulator (Leskiw et al., 1991a; Kwak et al., 1996). In-frame TTA codons are predominantly found in genes encoding regulators of antibiotic biosynthesis and antibiotic resistance genes with little representation in structural or enzymatic genes (Leskiw et al., 1991a; Leskiw et al., 1991b). Sequence gazing of the completed *Streptomyces* genome sequence has identified TTA codons within genes involved in sporulation, extracellular enzymes and a variety of genes without a known function, with very few identified genes that are linked to hyphal morphogenesis (Chater, 1998). One example of a developmental gene that contains TTA codons is *bldH*, which suggests a degree of *bldA* dependence and explains the grouping of *bldA* and *bldH* within the same complementation group (Champness, 1988; Chater and Horinouchi, 2003). Further evidence of BldH dependence upon the *bldA* encoded tRNA comes from a recent study where *S. coelicolor bldH* was determined to be a close homologue of *adpA* in *S. griseus*. AdpA is of the AraC family of putative DNA-binding transcriptional regulators and in *S. griseus* it is known to be essential in streptomycin biosynthesis and aerial growth in part through activation of the *adsA* encoded sigma-factor. In *S. coelicolor*, *adpA* (=*bldH*)

contains a single TTA codon and mutagenesis to form the alternate leucine codon TTG, was shown to partially restore aerial mycelium formation to a *bldA* mutant. Most interestingly, complementation only occurred when the TTA-free *adpA* was coupled downstream to *ornA*, which is naturally TTA-free and located downstream of *adpA* in its normal genomic position. This suggests that expression from TTA codon containing genes may affect the regulation of downstream genes and accordingly, polar mutations may occur within TTA-free genes in *bldA* mutants that substantially contribute to the mutant phenotype. It was also determined that *S. coelicolor adpA* is not directly regulated by γ-butyrolactones or does it appear to effect the activation of the *bldN* encoded sigma-factor, the orthology of the *S. griseus asdA*-encoded sigma factor. Consequently, morphological developmental defects in *S. coelicolor bldA* mutants appear to be largely due to the *bldH* (*adpA*) and *ornA* dependence upon BldA leucyl-tRNA (Chater and Horinouchi, 2003). A variety of signaling molecules and membrane protein-encoding genes also contain TTA codons and may be linked to the production of the second putative signal in the *bld* gene extracellular signaling cascade in parallel with AdpA activation (Chater and Horinouchi, 2003). However, no TTA codons appear to reside in genes concerned with vegetative growth or primary metabolism, which substantiates the notion that *bldA* is involved exclusively in the regulation of development (Leskiw et al., 1991a; Leskiw et al., 1991b).

Insofar as antibiotic production is concerned, the role of *bldA* tRNA has been elucidated. UUA codons are present within the transcripts of pathway-specific activators of pigmented antibiotic production in *S. coelicolor* and the *bldA* encoded leucyl-tRNA has been shown to be essential for the efficient translation of ActIIORF4 and RedZ, the

pathway-specific regulators of actinorhodin and undecylprodigiosin biosynthesis, respectively (Bystrykh et al., 1996; White and Bibb, 1997; Guthrie et al., 1998).

The global regulatory features of the DNA binding BldD repressor protein evidently link the stress response, antibiotic production, aerial hyphae formation and sporulation within an integrated circuitry. Characterization of the BldD protein has described it as a small, highly charged potential transcriptional regulator which binds DNA as a dimer and has a potential receiver and transmitter domain separated by a central proline and glycine rich flexible tether (Elliot et al., 1998; Elliot and Leskiw, 1999; Elliot et al., 2003b). BldD binds as a dimer to an imperfect inverted repeat, which can be numerously represented in the promoter regions of the genes it controls. Unlike other DNA binding proteins, BldD binding half sites are separated by variable spacing, which may alter the strength of BldD binding to different promoter regions. It is believed that the central linker region may allow a degree of flexibility in dimerization and binding to numerous *bldD* sites with inherent variable spacing (Elliot and Leskiw, 1999; Elliot et al., 2001). Futhermore, *bldD* is expressed at maximum abundance at early time points in vegetative growth and is postulated to repress developmental genes during vegetative growth and restrict premature morphological and physiological differentiation. The fact that *bldD* disruption imparts a *bld* mutant phenotype, with pleiotropic defects in both antibiotic production and aerial mycelium formation, suggests that the lack of the repressor may impair development due to the premature release and expression of developmental genes, or alternatively, may imply some unidentified activator function of BldD (Elliot et al., 1998).

Confirmed targets of direct BldD regulation include *bdtA*, encoding a putative transcription factor, and three σ factors, *sigH*, *whiG* and *bldN* (Elliot et al., 2001;

Kelemen et al., 2001). *sigH* encodes a stress-responsive σ factor and is found in an operon with a putative anti-σ factor gene. BldD binding represses expression from one of two (or four, in a separate publication (Kormanec et al., 2000)) *sigH* promoters (Kormanec et al., 2000; Kelemen et al., 2001). Upon BldD de-repression, sigHp2 is expressed specifically in aerial hyphae and marks *bldD* as a determinant of cell-type specific gene expression. Notably, this is the first direct link between stress and development and suggests that the stress response is an intrinsic feature of developmental control and that these regulatory networks overlap and interact (Kelemen et al., 2001). BldD also regulates *bldN*, a σ-factor involved in morphological differentiation but not antibiotic production, and *whiG*, a σ-factor which poises aerial hyphae for sporulation events (Elliot et al., 2001). This indicates a key role for *bldD* as an indirect regulator of differential expression from various subsets of genes. Interestingly, *ramR* and *ramC* appear to be indirectly regulated by the *bldD*- encoded repressor protein; therefore, BldD de-repression, in response to signals from the *bld* gene regulatory network, may link developmental transitions through coordination of *ram* gene induced aerial hyphae production with the *whiG*-directed initial priming events of sporulation (Chater, 2001; Elliot et al., 2001; O'Connor et al., 2002). However, a *bldD* connection to antibiotic production and the phenomenon of BldD de-repression is still unclear (Willey et al., 1993; Chater, 2001).

A direct developmental link to physiological stress was first glimpsed with the observation of *sigH* repression through BldD binding (Kelemen et al., 2001). Competing *sigH* studies have linked the σ-factor to the osmotic stress response, as well as heat and ethanol shock, and one study has revealed a key role for *sigH* encoded σ-factor in spore morphogenesis (Kormanec et al., 2000; Kelemen et al., 2001; Sevcikova et al., 2001).

Another σ-factor, *sigB*, has been observed to be involved in osmotic stress and development in *S. coelicolor*; thus, the products of *sigB* activation may aid the erection of aerial hyphae in a non-aqueous environment (Cho et al., 2001). Recently, studies have identified nine *sigB* homologs that may contribute to fine-tuning and control of the relative activity of developmental and stress-related promoters (Viollier et al., 2003). Other findings indicate that the oxidative stress response σ-factor, SigR, also plays a role in development (Paget et al., 1998; Kang et al., 1999). The interplay of multiple stress-responsive σ-factors and their paralogs with developmental genes could potentially act at various checkpoints preceding differentiation and may also be required by the mycelium to cope with various cellular stresses prompted by morphogenesis (Chater, 1998).

As stated above, a direct target of BldD repression during vegetative growth is *bldN*, encoding an ECF-like sigma factor, which is restricted to an involvement in morphological differentiation while not participating in the control of antibiotic production (Bibb et al., 2000; Elliot et al., 2001). *bldN* was initially identified as a *whi* mutant, with NTG-induced point mutations that reduced the activity or expression of σ^{BldN}. Strains with *bldN* mutant alleles either produced aberrant, elongated spores or long, undifferentiated aerial hyphae with rare, irregularly septated, spore chains. Although subsequent disruption of *bldN* yielded a *bld* phenotype, the phenotype of the *whi*-like *bldN* mutant alleles suggests that σ^{BldN} may also play a role in sporulation and that there is not a true delineation between the *bld* and *whi* genes as separate regulatory pathways controlling development. *bldN* does not fit into the speculative *bld* gene signaling hierarchy, for unknown reasons. *bldN* mutants restore fringe aerial hyphae production to *bldJ*, *K* and *H* but not to *bldA* or the other three complementation groups, and no *bld* mutants restore aerial hyphae production to *bldN*. Further characterization of

bldN revealed an indirect or direct dependence upon *bldH* and *bldG* for expression, as well as, direct repression by BldD binding at early vegetative time points (Bibb et al., 2000). σ^{BldN} is initially expressed as a preprotein, $\sigma^{proBldN}$, with an unusual 86 amino acid N-terminal extension, which potentially restricts σ-factor activity (Bibb and Buttner, 2003). The BldN preprotein processing agent has yet to be identified and may process the BldN pro-form in order to coordinate σ^{BldN} activation with other developmental events (Elliot et al., 2003b). The only direct target identified so far for active σ^{BldN} is *bldM*, encoding a putative response regulator; a direct dependence upon σ^{BldN} is still unreported for a variety of other *bldN* regulated genes that have been identified through microarray analysis, such as the chaplin genes (Bibb et al., 2000; Elliot et al., 2003a).

In vitro transcription studies determined that bldMp1, the first of two promoters which regulate *bldM* expression, is a direct target for σ^{BldN} activation (Bibb et al., 2000). BldM occurs within the same extracellular complementation group as *bldD* and encodes a response regulator of the FixJ subfamily, thereby, consisting of a C-terminal helix-turn-helix DNA binding motif and a common N-terminal phosphorylation pocket. Although the phosphorylation pocket is highly conserved among other response regulators, the putative phosphorylation site is not required for the function of BldM. Interestingly, *bldM* was identified via NTG-mutagenesis screens with the corresponding mutants described as *whi* mutants, with separate mutant alleles displaying a variety of sporulation-deficient phenotypes. Later, complete disruption of *bldM* yielded a *bld* phenotype. Consequently, different mutant alleles of *bldM* affect distinct stages in development, which is similar to mutant alleles of *bldN*, and suggests that *bld* genes that function late in development, as defined by the *bld* gene signaling hierarchy and *bldD* dependence, may have far-reaching functions past initial morphogenesis and into the dominion of

sporulation events. BldM mutant alleles exhibiting *whi* phenotypes contained point mutations in sequences coding for the well-conserved phosphorylation pocket. Inexplicably, the phosphorylation site was found to be dispensible, while the integrity of the entire pocket was critical for BldM function. Speculation persists that the phosphorylation pocket could possibly be important for interaction with a co-regulator of BldM function (Molle and Buttner, 2000). Furthermore, no direct BldM targets have been determined, although indirect *bldN*-regulated induction of the family of hydrophobic chaplin proteins, which are partially responsible for the fibrous sheath surrounding the aerial hyphae and spores, prompted binding studies of BldM with *chp* promoter fragments. No direct interaction was revealed, however this may be a somewhat inconclusive finding as a BldM-bound positive control has not been established, and the potential for the *in vitro* absence of a co-regulator may be significant (Elliot et al., 2003a).

A reasonable question that can be asked of *Streptomyces* research concerns the failure to identify structural genes with a *bld* mutant phenotype as targets of *bld* gene-encoded regulatory proteins. Shouldn't some key structural or enzymatic genes be imperative to morphological differentiation and thereby be identified in screens for *bld* mutants? A partial solution to this conundrum resides in the observation of functional redundancy for developmentally related physiological adaptations, where there are many structural elements which play overlapping functions in the achievement of development and morphogenesis (Chater, 1998). For instance, a study of the newly identified *chp* (chaplin) genes, which pertain to the hydrophobic fibrous sheath necessary for aerial hyphae erection and spore formation, revealed at least eight paralogous *chp* genes. Disruption mutagenesis of single *chp* genes had little discernible phenotypic effect; only

in combinations of quadruple and quintuple knockout mutants were deficiencies in differentiation observed (Elliot et al., 2003a). The redundant gene products may promote minute structural differences that cannot be identified by examining gross colony phenotype or respond to separate environmentally induced regulatory factors. Only when key regulators of entire subsets of genes, such as the *bld* genes, are disrupted, do severe developmental defects occur (Chater, 1998).

The *bldG* locus is of particular interest to my own research project and consists of *bldG* and *orf3*, which encode a putative anti-anti-σ factor and an anti-σ factor, respectively. Predicted BldG and ORF3 proteins show similarity to the *B. subtilis* RsbV and SpoIIAA anti-anti-σ factors, and RsbW and SpoIIAB anti-σ factors, respectively. Interestingly, RsbV/W are stress-responsive and SpoIIAA/AB play a role in sporulation-specific gene expression. In *B. subtilis*, anti-anti-σ factors, anti-σ factors and a cognate σ factor are expressed as polycistronic transcripts. The σ factor is inactivated through binding and sequestration by the cognate anti-σ factor. Anti-anti-σ factor alleviation of σ factor inhibition is dependent upon the phosphorylation state of a conserved serine residue within the anti-anti-σ factor. Moreover, phosphorylation by kinase domains of the related anti-σ factor, effectively inactivates the anti-anti-σ factor and prevents it from impairing anti-σ factor repression of the σ factor. Activation of a phosphatase occurs during periods of reduced ATP energy levels, for the Rsb system, or upon response to certain environmental cues, for the SpoIIA system. Concurrent dephosphorylation of the anti-anti-σ allows it to bind the σ factor bound anti-σ, thereby, mediating a conformational change in the anti-σ factor and allowing release of an active σ, which can now form a holoenzyme complex with RNA polymerase and direct specific gene expression. The *bldG* locus of *S. coelicolor* is loosely based upon this paradigm but does

have some significant divergence from *B. subtilis* σ factor regulation (Bignell et al., 2000). For instance, the BldG protein contains the conserved serine phosphorylation site, which is inherent in this class of anti-anti-σ factors, and phosphorylation appears to occur *in vitro* and *in vivo* (Bignell et al., 2000; Bignell et al., 2003). However, *in vitro* phosphorylation was not mediated by the *orf3* encoded anti-σ factor, which lacks three of five conserved amino acid residues that are important to kinase activity. It is possible that a co-factor or a separate kinase is responsible for BldG phosphorylation and presumable inactivation (Bignell et al., 2003). In addition, besides being transcribed as a polycistronic transcript containing *bldG* and *orf3*, *bldG* is also expressed as a monocistronic transcript. Both transcripts are present at all growth periods, yet the monocistronic transcript is two to three times more abundant and both are upregulated at periods corresponding to aerial hyphae formation and antibiotic production. Perhaps the most significant difference between the *bldG* locus and similar regulatory pairs in *Bacillus* is the lack of a genomic link to a cognate σ factor. It has been postulated that BldG and ORF3 may regulate the activity of a number of σ factors or direct the activity of a single σ factor only under a subset of environmental conditions (Bignell et al., 2000). It is also possible that separate *bldG*-controlled σ factors are responsible for hyphal morphogenesis and antibiotic production.

In *Bacillus*, regulatory pairs like SpoIIAA/AB and RsbV/W govern a subset of σ^{70}- like σ factors, and there are at least nine *S. coelicolor* homologs within this subfamily (Bignell et al., 2000). Notably, disruption of the *bldG* gene yields a *bld* phenotype; whereas, *orf3* disruptions have an indeterminant phenotype which may only be viable due to second site mutations; therefore, σ factors regulated by the *bldG* locus are directly involved in development and may have toxic effects if released prematurely

(Bignell et al., 2000; Bignell et al., 2003). Although having no genomic link to a cognate σ factor, the *bldG* locus is positioned closely divergent to a gene, designated *sco3550* (Stoehr, 2001), encoding a putative RNA helicase (Bignell et al., 2000). It appears probable that overlapping promoters of the *bldG* locus and *sco3550* locus lead to coordinated expression of the loci through some fashion of RNA polymerase holoenzyme steric hindrance or a supercoiling dependent mechanism (Stoehr, 2001). My own MSc project will attempt to assess the effect of this putative RNA helicase upon the *bldG* locus and provide an evaluation of the interplay of the helicase in the *bld* gene regulatory network and in *S. coelicolor* development (see section 4.0.).

1.6 RNA Helicases

RNA helicases are classically defined as enzymes which utilize NTP hydrolysis to facilitate double-stranded RNA unwinding or duplex destabilization. Criteria for RNA helicase characterization usually entail the presence of RNA-dependent NTP hydrolysis activity (primarily ATP hydrolysis) coupled to either 5'-3' or 3'-5', ATP- dependent RNA unwinding (Gorbalenya et al., 1989). Predicted RNA destabilizing enzymes are ubiquitous in all three domains of life, as well as the viruses, and are involved in nearly all aspects of RNA metabolism. Furthermore, RNA helicases play highly specific roles in viral propagation, RNA degradation, transcription, translation initiation and termination, mRNA localization, ribosome biogenesis, mRNA editing, pre-mRNA splicing, RNA interference and RNA nuclear export. Evidence of high sequence and organizational conservation of RNA helicase motifs, combined with the diverse function and abundance of these enzymes in nature, suggests that RNA helicases are ancient enzymes whose progenitors would have represented a keystone step in the evolution of the early RNA world (Tanner and Linder, 2001a).

1.7 The RNA Folding Problem

It is now generally accepted that many RNA molecules require a biologically active form in order to interact with other nucleic acid and protein macromolecules and therefore must fold into specific intramolecular structures. Fundamental folding problems inherent in single-stranded RNA molecules contribute to the cellular necessity for RNA chaperones, such as non-specific and specific RNA binding proteins, and RNA helicases. The major RNA folding problem arises from the lack of significant structural and chemical differences between the four nucleotides contained within RNA. Consequently, single stranded RNA need only find a small complementary sequence to conduct base-pair formation and thus, a myriad of alternate RNA secondary structures with similar thermodynamic stabilities can be formed. This contrasts with more defined protein folding, which forms fewer alternate conformations than that of RNA due to the potential combinations of 20 amino acids with varying structures and hydrophobic, hydrophilic or charged properties. Alternate RNA tertiary structures can also arise because RNA secondary structure does not present enough information to influence favorable folding. An absence of sequence information in RNA secondary structure is due to the inwards orientation of base pair interactions; therefore, secondary structures containing different primary nucleotide sequences appear structurally similar and it is difficult for proper folding to generate a unique tertiary structure. The ease of formation of alternate RNA conformations is detrimental to cellular processes as a functional RNA molecule often has a specific tertiary structure. Also, alternative RNA structures may have a higher or similar thermodynamic stability than that of active RNA structures and are effectively kinetically trapped in these configurations (Herschlag, 1995). RNA chaperones become obligatory in these circumstances to resolve misfolded RNA

structures and allow proper folding to occur (Herschlag, 1995; Lorsch, 2002; Mohr et al., 2002). The principle behind chaperone activity pertains to a lowering of transition barrier energy to allow functionally unfavorable RNA to unfold. Lowering of unfolding kinetic energy is most likely accomplished through complementary charge interactions between the RNA and the association of non-specific RNA binding proteins, or by ATP hydrolysis driven RNA destabilization by RNA helicases (Herschlag, 1995). Conversely, RNA chaperones may bind to single stranded RNA as a preemptive measure to prevent unfavorable intramolecular RNA structures from forming. Passive refolding of chaperone unfolded RNA is usually restricted since unwound RNA would most likely form an alternate conformation, especially if the active RNA conformation does not invoke maximum thermodynamic stability. Rather, continued surveillance of the RNA folding state by RNA chaperones is aided by the action of RNA- specific binding proteins that interact with the functional RNA conformation and stabilize it in an active state (Herschlag, 1995; Mohr et al., 2002).

It is difficult to assess the RNA chaperone activity of RNA helicases *in vivo*, as RNA unfolding and refolding is a dynamic and rapid process, and the extent to which RNA helicases participate in cellular RNA chaperone activity is still unclear. To date, only a *Neurospora crassa* RNA helicase, CYT-18, involved in spliceosome pre-mRNA splicing, has been implicated in a cellular RNA chaperone role (Lorsch, 2002; Mohr et al., 2002). While at least four RNA helicases have been observed to play a more active role in RNA structural rearrangements and possess both RNA unwinding and intrinsic ATP-independent RNA annealing activity. For instance, yeast RNA helicases, p68 and p72 and cyanobacterial RNA helicase, CrhC, are believed to anneal complementary ssRNA into dsRNA and are believed to catalyze strand exchange and annealing of

complementary RNA into intermolecular duplexes (Rossler et al., 2001; Chamot et al., 2004). Interestingly, biochemical evidence shows that the majority of RNA helicases have little processivity, except for viral helicases, and are commonly residents of complex macromolecular assemblies (von Hippel and Delagoutte, 2001). These findings have precipitated the belief that RNA helicases have highly specific cellular roles to govern the spatial and temporal progression of dynamic biochemical exchanges in biological processes involving RNA (Tanner and Linder, 2001a). For instance, at least 17 distinct RNA helicases are implicated in yeast ribosome biogenesis and most likely moderate RNA unfolding and passage to a complementary RNA or a specific RNA binding protein in the context of ribosome machinery construction (Kressler et al., 1999; Tanner and Linder, 2001b).

1.8 RNA Helicase Taxonomy and the Structure and Function of Conserved Motifs

Based upon comparative analysis of sequence similarity and the general organization of conserved motifs, RNA (and DNA) helicases are systematically categorized into five superfamilies (SF), while position in a superfamily does not correlate to nucleic acid substrate preference (Gorbalenya and Koonin, 1993). SF1 and SF2 have eight to nine common motifs separated by poorly conserved spacer regions. SF1 primarily contains DNA helicases, such as *Escherichi coli* UvrD-like and Rep-like helicases, as well as some viral RNA helicases; whereas most RNA helicases have been categorized as members of SF2, which all contain variations upon a conserved D-E-A-D (Asp-Glu-Ala-Asp) motif (Linder et al., 1989; Pause and Sonenberg, 1992; Gorbalenya and Koonin, 1993). This conserved motif may vary indefinitely at the positioning of the alanine residue while, in some RNA helicases, the final aspartic acid residue is replaced

by a histidine residue; therefore, three subfamilies of RNA helicases, containing DEAD, DEAH or DExH motifs, are found within SF2 (Pause and Sonenberg, 1992; Schmid and Linder, 1992). The sequence characterization of newly identified helicases and the solved crystal structure of representative members of each superfamily has made it is increasingly clear that SF1 and SF2 are closely related structurally and would be more appropriately recognized as separate subsets of the same family (Korolev et al., 1998). Also, SF1 and SF2 helicases function as monomers or dimers, whereas, SF3, SF4 and SF5 contain divergent hexameric helicases with fewer conserved motifs, such as the RNA helicase required for Rho-dependent transcriptional termination (Gorbalenya and Koonin, 1993; Tanner and Linder, 2001b).

Helicases of SF1 and SF2 have eight to nine common motifs that are widely dispersed in primary sequence but are clustered within a tertiary structure to form a conserved catalytic core that is essential for a common mechanism of NTP hydrolysis driven nucleic acid destabilization (Gorbalenya and Koonin, 1993; Caruthers and McKay, 2002). Site-directed mutagenesis studies of structure-function relationships of motifs indicate that all motifs are required for the cellular helicase enzymatic mechanism. Although overlapping functions do exist, conserved motifs primarily regulate NTP binding and hydrolysis, oligonucleotide substrate binding, or the coupling of NTP hydrolysis to domain movement and RNA unwinding (Hall and Matson, 1999; Caruthers and McKay, 2002). Helicase motifs are denoted by their positions relative to motifs I and II which are the Walker A and Walker B motifs, respectively. These are well- conserved NTP binding motifs that are common to all superfamilies of helicases, as well as NTP-binding proteins with NTPase activity (Walker et al., 1982). The Walker A motif, described as a GxGKT/S consensus sequence, specifically binds the phosphates of a

bound NTP, while the Walker B motifs, described by DExH/D (in SF2) consensus sequences, binds NTP phosphates through the coordination of a Mg^{2+} ion and effect NTP hydrolysis by the addition of a water molecule (Walker et al., 1982; Pause and Sonenberg, 1992). Motifs Ia and IV specifically interact with nucleotide substrates and commonly impart binding through interactions with the sugar phosphate backbone (Kim et al., 1997b; Korolev et al., 1998; Lin and Kim, 1999; Caruthers et al., 2000). Finally, motif III, defined as a SAT consensus sequence, appears to couple Walker domain movement upon NTP binding and hydrolysis to enzyme translocation and RNA destabilization (Gorbalenya et al., 1989; Pause and Sonenberg, 1992; Schwer and Meszaros, 2000). Motifs V and VI also appear to play a role in both oligonucleotide and NTP binding and may be involved in the transduction of energy yielded from NTP hydrolysis to domain movement and RNA destabilization (Pause et al., 1993; Gross and Shuman, 1996). The mechanistic role of transducer motifs III, V and VI is poorly defined and there appears to be fairly high sequence variation between different subsets of helicases (Figure 1.3 A, B).

Structure-function analysis of SF1 and SF2 conserved helicase motifs defines a central core that is essential for the helicase biochemical mechanism (Kim et al., 1997a; Caruthers et al., 2000). Conversely, substrate specificity is believed to be invoked by long N or C terminal peptide extensions, the interaction of helicases with accessory factors, or the integration of an RNA helicase within a specific macromolecular complex, thus yielding a specific cellular context for RNA unraveling (Tanner and Linder, 2001a). The dangers of assigning RNA helicase activity based upon comparative sequence analysis alone are exhibited by the Snf/Swi family of proteins from yeast and higher eukaryotes. These proteins contain the eight conserved motifs characteristic of helicase

activity, but function in chromatin remodeling, and in part, transcription initiation rather than double-stranded nucleic acid unwinding (Havas et al.). Consequently, the helicase catalytic core is more broadly viewed as a non-partisan engine which promotes nucleic acid structural remodeling, while the distinct function of a helicase is defined by unconserved stretches of amino acids within the enzyme. These variable domains may specify helicase activity by positioning the active helicase within a specific cellular context through protein-protein interactions, oligomerization regulation and substrate specificity.

A common problem previously encountered when comparing site-directed mutagenesis structure-function studies between subfamilies of RNA helicases in different organisms was the lack of knowledge of specific biological roles for these enzymes. RNA helicase RNA chaperone activity was an unproven hypothesis that had only been observed within the confines of *in vitro* experiments and has only recently been described *in vivo*, in *N. crassa* (Mohr et al., 2002). A description of the crystal structures of SF2 helicase eukaryotic initiation factor 4A (eIF4A), hepatitis C virus (HCV) non-structural protein 3 (NS3), and a DEAD box helicase from the archaebacterium, *Methanococcus jannaschii* (MjDEAD), has allowed structural comparisons with SF1 helicases (Yao et al., 1997; Cho et al., 1998; Kang et al., 1998; Caruthers et al., 2000; Story et al., 2001). For the most part, the central catalytic cores of SF1 and SF2 helicases can be structurally superimposed and this high structural conservation points to a degree of functional conservation, which has alleviated concerns of the validity of structure-function comparisons between different subsets of helicases (Yao et al., 1997; Cho et al., 1998; Korolev et al., 1998; Caruthers et al., 2000). The enzymatic core consists of two domains

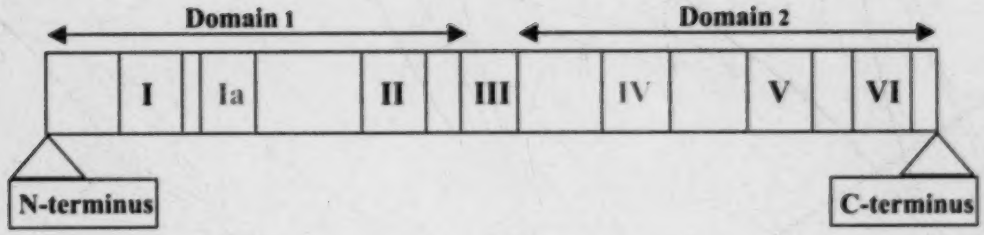
Figure 1.3 Structural Organization of Conserved Helicase Motifs

(a) General organization of conserved helicase motifs subdivided into two domains. The domains include motifs which are believed to function in NTP hydrolysis (black), non-specific substrate binding (red) and motifs that link these two activities to substrate unwinding (blue). Adapted from (Tanner et al. 2001).

(b) Motifs I-IV consensus sequences for the three subfamilies of RNA helicases, including the putative *S. coelicolor* DECH box RNA Helicase, SCO3550. Note that amino acid spacing between motifs of SCO3550, which is representative of other RNA helicases, indicates a clustered catalytic core. In SCO3550, a large C-terminal extension is also seen which may contribute to substrate specificity. Adapted from (Tanner et al. 2001; Stoehr 2001).

(c) The predicted crystal structure of *M. jannaschii* DEAD box RNA helicase (MjDEAD) reveals a small protein where all conserved motifs face inwards and potentially have overlapping functions. A tether is shown to separate the two domains and creates a degree of flexibility for enzymatic conformational changes. Adapted from (Story et al. 2001).

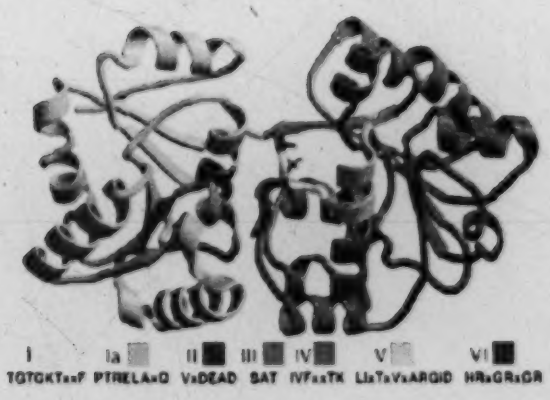
(A)



(B)

	Walker A Motif I	Motif Ia	Walker B Motif II	Motif III	Motif IV	Motif V	Motif VI
DEAD	ATGTGKT	—PTRELa-Q	VlDEaD-m	SAT	LiF—T	lvaTDvaaRGlD	Y-HRiGRTgR-G
DEAH	GETGTGKS	-TQPRR-aA	i-DEaHER	SAT	LvFL-G	TNIAETS-Ti-g	a-QR-GRAGR—
SCO3550	GKS	PTKALA	DECH	SAT	LVFAL	LAGRVAAYRG	QQAGRAGR
	76aa	29aa	72aa	33aa	28aa	65aa	52aa 404aa

(C)



connected by a flexible tether and allows space for NTP binding. All conserved motifs face into the NTP binding cleft; therefore, motifs cannot be easily categorized as possessing a singular, specific function, and all conserved motifs appear to have a hand in NTP binding, hydrolysis and coupling this energy to enzyme conformational changes and disruption of RNA secondary structure (Cho et al., 1998; Story et al., 2001). There are also several isolated conserved amino acids that directly interact with nucleotide bases, and subsequent NTP hydrolysis may alter the spacing of these amino acids and thus, serve to unravel RNA secondary structure [Figure 1.3 C]. However, the same caveat applies that variable sequence outside the conserved core may be responsible for the direction of the specific biological role of the helicase and may modify the catalytic mechanism and efficiency (Hall and Matson, 1999; Caruthers and McKay, 2002).

Assuming no interference from variable peptide sequences at the periphery of helicase conserved motifs, the conservation of SF1 and SF2 catalytic core structures lends support to three separate theories of active RNA unwinding based upon the oligomerization state of the enzyme. The inchworm model predicts the action of a RNA helicase monomer and is accepted as the most consistent model with current biochemical evidence. This model states that ATP binding and hydrolysis effects conformational changes between the two domains of the catalytic core. As the helicase structure changes, translocation of the enzyme is mediated by the dissociation and subsequent re-association of individual conserved amino acids that intercalate with the nucleotide substrate. Release of ADP and further binding of ATP poises the helicase downstream of the previous round of ATP hydrolysis and unwinds RNA secondary structure through disruptive translocation (Velankar et al., 1999). The active rolling model presumes a dimeric helicase structure where each subunit assumes a separate conformation upon

ATP binding and hydrolysis. One subunit would have specificity for single stranded RNA and the other subunit for double stranded RNA. These conformations would shift due to ATP hydrolysis and drive a see-saw-like translocation to disrupt RNA secondary structure (Cho et al., 1998; Lin and Kim, 1999). Lastly, hexameric helicases of SF3, SF4, and SF5 are believed to swivel around RNA and thread single-stranded RNA through the open core of the doughnut-like arrangement with the complementary strand being threaded around the outside of the helicase (Hall and Matson, 1999). Each of these models may be correct for certain biological processes that require a defined factor of helicase processivity and efficiency. Although all three models predict translocation mediated secondary structure denaturation, biochemical studies have shown that most RNA helicases require single stranded RNA loading, have little processivity, and have putative substrates with small regions (<10bp) of secondary structure; therefore, RNA helicases have been described as a functional clamp to prevent RNA misfolding and mediate progressive RNA-RNA associations (Tanner and Linder, 2001a).

1.9 The Action of RNA Helicases in Nature- a Broadening Definition

RNA helicases have been described in all classes of organisms and implicated in numerous aspects of RNA metabolism. While many newly identified putative RNA helicases have the characteristic core motifs that specify general nucleic acid structuring capability, more in-depth genetic and biochemical analyse have uncovered new enzymatic properties that expand upon established paradigms of helicase activity (Gorbalenya and Koonin, 1993; Tanner and Linder, 2001b). These newly defined characteristics often confirm *in vitro* hypotheses, such as RNA chaperone activity, as well as identify new cellular niches for enzymatic activity and clarify the biological role for RNA helicases in elaborate macromolecular mechanisms.

Studies of viral RNA helicases have confirmed unifying concepts found in the RNA helicase body of research and have also revealed novel RNA helicase properties, which may translate to subsets of prokaryotic and eukaryotic helicases. Much of the primary data for viral RNA helicases have come from studies on the Hepatitis C viral DExH box RNA helicase, non-structural protein 3 (NS3), and a *Vaccinia* viral DExH RNA helicase, nucleoside triphospate phosphohydrolase II (NPH-II) (Shuman, 1992; Jin and Peterson, 1995). The high resolution crystal structure has been solved for HCV NS3 and has contributed to the establishment of the active rolling model of dimeric helicase activity. NS3 is a Y-shaped protein with three domains, where the third domain acts as a linker between domains 1 and 2, which in turn contain motifs responsible for enzymatic catalysis; thus, subsequent dimerization would occur and conformational transitions between domains 1 and 2 would mediate RNA unwinding by an NTP switch (Cho et al., 1998). Interestingly, the entire HCV genome is expressed as a polyprotein that is proteolytically processed to form viral proteins, including NS3, which is a bifunctional protein with both a serine-protease domain and an RNA helicase domain. These domains should not be confused with the three domains observed by x-ray crystallography as these are all sub-domains of the NS3 RNA helicase domain, while the serine protease domain was removed to facilitate crystallization (Jin and Peterson, 1995; Cho et al., 1998). NS3 is believed to function in a yet untold fashion in cooperation with the NS4A cofactor and RNA-dependent RNA polymerase, NS5B, for viral genome replication, and may represent a viable clinical target for anti-viral treatment. Unique characteristics constituting NS3 activity include the ability to unwind RNA/RNA, RNA/DNA and DNA/DNA duplexes, as well as high processivity upon DNA templates, while processivity on RNA templates is dependent upon NS4A cofactor binding (Locatelli et

al., 2002; Pang et al., 2002). This is also the first example of a multifunctional protein containing a RNA helicase domain, and the first instance of functional diversity in an RNA helicase due to structural organization by a cofactor. Because HCV viral replication does not transfer through DNA intermediates, the RNA helicase cofactor-mediated activity may affect viral genome replication, whereas DNA helicase activity may alter the genome of the eukaryotic host cell (Pang et al., 2002).

The DExH RNA helicase, NPH-II, of infectious *Vaccinia* virus functions as a processive, unidirectional motor that is critical for viral replication and is believed to act within a transcriptional elongation complex with viral RNA polymerase to direct early mRNA expression of viral genes (Shuman, 1992; Jankowsky et al., 2000). An exciting study performed on NPH-II DExH/D box protein potentially redefines the role for certain RNA helicases as RNPases (ribonucleoprotein displacement upon NTP hydrolysis); enzymes which preferentially disrupt RNA-protein interactions rather than provide direct RNA chaperone activity to duplex RNA. The authors postulated that a novel role for certain RNA helicases is to foster the peptide structural reorganization of ribonucleoprotein assemblies, either by changing the RNA structure of the protein binding site or by forcing the protein off the RNA by processive translocation of the helicase. However, RNA helicase protein dissociation studies were only demonstrated with *in vitro* experiments and RNA-protein disruption may be a secondary artifact of high NPH-II processivity. Therefore, the physiological relevance of DExH/D box protein RNPase activity is disputed and may not indicate a cellular activity of NPH-II or correspond to novel activities of DExH/D box proteins of higher organisms (Jankowsky et al., 2001; Schwer, 2001).

The most well characterized eukaryotic RNA helicases are members of ribonucleoprotein assemblies that function in translation initiation, ribosome biogenesis and pre-mRNA splicing. The yeast RNA helicase, eukaryotic translation initiation factor 4A (eIF4A), acts in conjunction with other translation initiation factors to unwind secondary structure formation in the 5' untranslated regions of mRNA, in order to facilitate ribosome binding. The solved crystal structure and mutational studies of the 394 amino acid eIF4A indicate that the enzyme represents a minimum helicase enzymatic core, with little N- or C-terminal extension past the conserved helicase motifs and high structural conservation between the caltalytic sites of other SF1 an SF2 crystal structures (Pause and Sonenberg, 1992; Caruthers et al., 2000; Tanner and Linder, 2001b). eIF4A exhibits intrinsic weak helicase activity and requires association with the eIF4B cofactor to stimulate maximal enzymatic activity and novel bidirectional unwinding activity. eIF4A is also dependent upon interaction with other translation initiation factors for sequestration to mRNA 5' UTR secondary structure (Rozen et al., 1990; Pause and Sonenberg, 1992). Clearly, N- and C-terminal peptide extensions in other helicases, which are lacking in eIF4A, may act as non-conserved motifs which provide efficiency, processivity and substrate specificity to the enzyme, and are analogously represented by eIF4A cofactor binding.

An analysis of eukaryotic RNA helicases involved in ribosome biogenesis and pre-mRNA splicing presents informative models for the dynamic biological processes which involve RNA. Moreover, yeast ribosome biogenesis involves 78 ribosomal proteins, four rRNAs, numerous snoRNAs (small nucleolar guide RNAs) and at least 17 RNA helicases (Kressler et al., 1999). The finding that all nucleolar RNA helicases have highly specific activity and are essential for ribosome maturation, led researchers to

postulate that RNA helicases may mediate individual steps in the pathway and give directionality to the progression of a multi-step reaction (Colley et al., 2000). Multiple predicted RNA helicases mediating separate steps in a complex ribonucleoprotein complex are also apparent in the yeast spliceosome, which contains five snRNAs (small nuclear RNAs) and greater than 50 proteins, with at least eight of these DExD/H proteins (Schwer and Gross, 1998; Schwer and Meszaros, 2000). As previously mentioned, *N. crassa* RNA helicase, CYT-19, was shown to participate in group I intron pre-mRNA splicing in a RNA chaperone role. CYT-19 presumably works in conjunction with a specific RNA binding spliceosome protein, CYT-18, which primarily acts to stabilize group I intron structure in an active form. The RNA chaperone activity of CYT-19 is required to unfold introns that have misfolded into a stabilized, inactive form (Lorsch, 2002; Mohr et al., 2002). Finally, many RNA helicases have been implicated in disease processes in multi-cellular eukaryotes, as well as serving as important factors in development, particularly in the control of cellular polarity of gene expression, such as the DEAD box RNA helicase, VASA, which has human and mouse homologs and is best characterized for its role in *Drosophila* germ cell development (Lasko and Ashburner, 1988; Castrillon et al., 2000).

A multitude of RNA helicases exist in archaebacteria and eubacteria and, similar to helicases in eukaryotes and viruses, they have diverse biological functions pertaining to cellular RNA processes. An informative example of a bacterial RNA helicase involved within a macromolecular complex is that of *E. coli* RhlB, which plays an integral role in mRNA turnover and is a vital component of the RNA degradosome. mRNA degradation is an important determinant of proteome regulation within the cell, and mRNA turnover is believed to be primarily mediated by a multi-enzyme complex

containing the RNaseE endoribonuclease, the polynucleotide phosphorylase (PNPase) exoribonuclease, the RhlB DEAD box helicase, and, for unknown reasons, an enolase glycolytic enzyme (Miczak et al., 1996; Py et al., 1996; Regnier and Arraiano, 2000). Because the PNPase exoribonuclease removes terminal nucleotides in a 3'→5' processive direction, mRNA 3' secondary structure can greatly impede and even halt exonucleolytic activity. RnaseE endonucleolytic activity is also specific for only single stranded RNA and is susceptible to mRNA secondary structure alterations. Consequently, the RhlB RNA helicase is required to remove RNA secondary structure and thereby increase the processivity of the degradosome complex. Unlike RNA helicase chaperone activity or RNPase activity, the endpoint of RhlB duplex destabilization would be ssRNA degradation with no ensuing refolding or alternate conformation stabilization. RNA helicase interactions with macromolecular complexes that direct RNA decay are a common phenomenon in prokaroyotes and eukaryotes, and there have been several attempts, with varying degrees of success, to integrate RhlB-activated RNA degradation by PNPase into the active rolling model or inchworm model for RNA helicase unwinding activity (Py et al., 1996; Coburn et al., 1999).

An examination of most identified RNA helicases has yielded little data concerning *in vitro* substrate specificity. The association of protein factors to N- and C-terminal peptide extensions is believed to mediate substrate specificity, as well as the cellular context of the substrate, and thus *in vitro* experiments rarely simulate the complexity of macromolecular assemblies (Tanner and Linder, 2001a). However, substrate specificity has been observed for the *E. coli* RNA helicase, DbpA, and the *B. subtilis* homolog, YxiN. These enzymes are implicated in ribosome biogenesis, and *in vitro* experiments showed that DbpA and YxiN will bind a variety of substrates, but only

an interaction with hairpin 92 of 23S rRNA stimulates ATP hydrolysis and RNA helicase activity (Fuller-Pace et al., 1993; Nicol and Fuller-Pace, 1995; Boddeker et al., 1997; Kossen and Uhlenbeck, 1999). Recently, this specificity for 23S rRNA has been shown to be sufficiently generated by a conserved C-terminal domain of DbpA and YxiN (Kossen et al., 2002). Consequently, the catalytic core of RNA helicases is most likely responsible for generic, indiscriminate substrate binding, while variable N and C terminal extensions moderate substrate specificity in some helicases (Tanner and Linder, 2001b; Kossen et al., 2002).

2.0 Thesis Objectives

Given the abundance of RNA helicases in nature, it seems a natural extension to identify these enzymes in the streptomycetes. Furthermore, annotation of the *S. coelicolor* genome sequence has predicted at least 13 RNA helicase genes. As aforementioned, initial characterization of the *bldG* and *orf3* genes identified a closely linked and divergently transcribed partial open reading frame. The complete DNA sequence and corresponding amino acid sequence BLAST-P search revealed that the open reading frame, designated *sco3550*, showed sequence similarity with the SF2 helicases (Bignell et al., 2000). Further analysis revealed seven conserved RNA helicase motifs, including the Walker B DECH (DExH) box. Also, a large C-terminal extension is believed to be responsible for conferring RNA substrate specificity as previously discussed [Figure 1.3 B] (Stoehr, 2001).

It was the goal of previous MSc student Julie Stoehr to assess the developmental relevance of the putative RNA helicase through determination of temporal gene expression and the construction of a *sco3550* null mutant with accompanying phenotypic analysis. It was hoped that this would provide initial evidence of the interplay of the

putative helicase with the *bldG* operon and substantiate the role of the gene in the complex process of physiological and morphological differentiation in *S. coelicolor*. Possibly owing to 5' RNA secondary structure, the instability of the *sco3550* transcript or extremely low inherent gene expression, most conventional methods of measuring gene expression, such as northern analysis, promoter probe assays, S1 nuclease mapping and primer extension, proved fruitless. The helicase gene expression was only detected by reverse-transcriptase-polymerase chain reaction (RT-PCR) and semi-quantitative RT-PCR revealed a biphasic expression pattern with peaks at 15 hour and 36 hour time points corresponding to vegetative growth and aerial hyphae formation. RT-PCR also aided in the general localization of the SCH5.13 transcriptional start site and prompted a mapping shift away from the translation start site predicted by the *S. coelicolor* genome database to a new downstream ATG start site. Subsequent attempts to disrupt the helicase gene were unsuccessful, leading to the belief that the gene might be essential to *S. coelicolor* viability; however, a disruption of the helicase gene was constructed with relative ease in a strain with a *bldG* disruption. The tentative conclusion reached from this thesis project stated that the *sco3550* gene product may be essential to the organism by acting as a suppressor of sporulation during vegetative growth, where early activation of the differentiation genes would be fatal. A *bldG/sco3550* double mutant was predicted to be viable due to the absence of these detrimental differentiation genes in a *bldG* mutant background. Another possibility raised was that the gene product plays separate roles in the vegetative and sporulation states of the organism, as evidenced by the developmentally-related biphasic gene expression profile (Stoehr, 2001).

To further elucidate the role of the putative RNA helicase in both vegetative growth and the growth phase transition involving aerial hyphae, antibiotic production and

subsequent sporulation, my MSc project was aimed at further assessing both the gene and protein expression of *sco3550* and at making new attempts at mutagenesis. Due to the uncertainty in the prediction of function based upon sequence analysis alone, the predicted RNA unwinding activity of the enzyme was also validated. As well, it was of interest to determine if SCO3550 is indeed associated in a macromolecular complex and whether it functions in a general cellular RNA process, or a more defined developmental role where it may regulate a subset of genes and perhaps play a hand in the *bld* gene or *whi* gene regulatory network.

Chapter 2:

Materials and Methods

2. Materials and Methods

2.1.1 Bacterial Strains, Plasmids and Growth Conditions

2.1.1 *Streptomyces coelicolor* and *Escherichia coli* strains

E. coli strains used in this study are listed in Table 2.1. *S. coelicolor* strains used in this study are listed in Table 2.2.

2.1.2 Plasmid Vectors

Cloning vectors used in this study are listed in Table 2.3. Recombinant plasmids from this study, together with the cloning strategies used for their construction, are listed in Table 2.4.

2.1.3 Growth and maintenance of *E. coli* strains

Liquid cultures of *E. coli* strains were grown at 30 or 37 ° C, where indicated, in LB medium (5 g/L yeast extract, 10 g/L tryptone, 10 g/L NaCl) (Sambrook et al., 1989) on a rotating rack. Solid cultures of *E. coli* strains were grown on LB medium containing 1.5% w/v agar. When necessary, the medium was amended with various antibiotics including, ampicillin (100 μg/mL, Sigma), apramycin (50 μg/mL, Provel), chloramphenicol (25 μg/mL, Aldrich), kanamycin (30 μg/mL, sigma), spectinomycin (100 μg/mL, Sigma), and/or tetracycline (2.5-100 μg/mL, Sigma). Strains were maintained as frozen stocks in 20% glycerol (v/v) at -85°C.

Table 2.1: ***Escherichia coli*** **strains used in this study**

E. coli **strain**	**Genotype**	**Reference or Source**
DH5α	F', ϕ80*lacZ*ΔM15Δ(*lacZYA-argF*)U169, *deoR,recA1,hsdR17*(r_{k-},m_{k}) *phoA, supE44,*λ-, *thi-1,gyrA96, relA1*	(Hanahan, 1983); Gibco BRL
ET12567	F', dam13::Tn9, *dcm6, hsdR, recF143, zjj202::* TN10, *galK2, galT22, ara14, lacY1, xyl5, leuB6, thi1, tonA31, rpsL136, hisG4, tsx78, mtl1, glnV44*	(MacNeil, et al., 1992); Gift from D. MacNeil, Merck Sharp and Dohme Research Laboratories
BL21(DE3)pLysS	F', *ompT*, *hsdS*$_{B}$(r_{B-},m_{B-}), *gal, dcm,* (DE3), pLysS (Cm^{R})	Stratagene; Gift from Gary Ritzel, University of Alberta
BW25113	Δ*(araD-araB)567,* Δ*lacZ4787(::rrnB-4), lacIp-4000(lacIq),* λ-, *rpoS369*(Am), *rph-1,* Δ*(rhaD-rhaB)568, hsdR514*	(Datsenko and Wanner, 2000); Plant Bioscience Limited, U.K.

Table 2.2: *Streptomyces* strains used in this study

***Streptomyces coelicolor* A3(2)**	**Genotype**	**Reference source**
J1501	*hisA1, uraA1, strA1, pgl*, SCP1$^-$, SCP2$^-$	(Chater, et al. 1982);John Innes Institute
M145	prototrophic, SCP1$^-$ SCP2$^-$ Pgl$^+$	(Lomovskaya, et al. 1980); John Innes Institute
M600	SCP1$^-$, SCP2$^-$	(Chakraburtty and Bibb 1997); Gift from M. Buttner, John Innes Centre
*Δ bldG*3B/ Ap^R*sco3550*	M145 derivative with a disruption of *bldG* and a putative disruption of *sco3550*	(Stoehr 2001)

Table 2.3: Plasmids and cloning vectors used in this study

Plasmid	Antibiotic Resistance Marker	Relevant Characterisitcs	Reference (or Source)
***E. coli* Plasmids, Phagemids and Cosmids**			
pBluescript®KS	Ampicillin	High copy number phagemid containing T7 and T3 polymerase promoters	(Alting-Mees and Short 1989); Stratagene
pMAL-c2X	Ampicillin	pBR322-derived low copy number vector for overexpression of MBP-fusion proteins	New England Biolabs
pUZ8002	Kanamycin	RP4 plasmid derivative encoding products necessary to confer mobilization of *oriT*-containing plasmids from *E.coli* to *S. coelicolor*	Gift from M. Buttner, John Innes Centre
pIJ773	Ampicillin, Apramycin	pBluescript KS+ derivative containing the apramycin resistance gene [*aac(3)IV*], and *oriT,* flanked by FRT sites	(Gust, et al. 2003); Plant Bioscience Limited
pIJ790	Chloramphenicol	λRED recombination plasmid encoding *gam, bet* and *exo* under control of the arabinose-inducible p*ara*BAD promoter. Also contains the repA101 temperature-sensitive replicon	(Gust, et al. 2003); Plant Bioscience Limited
Cosmid SCH5	Ampicillin, Kanamycin	Supercos-1 derivative containing 40.54 kb of *S. coelicolor* chromosomal DNA, including the *bldG* operon and *sco3550*	(Redenbach, et al. 1996);Gift

Table 2.3 continued: Plasmids and cloning vectors used in this study

Plasmid	Antibiotic Resistance Marker	Relevant Characterisitcs	Reference (or Source)
***E. coli-Streptomyces* shuttle vectors**			
pSET152	Apramycin	High copy cloning vector in *E. coli*, integrates into the *Streptomyces attB* site	(Bierman, et al. 1992); Northern Regional Research Center, Peoria,Ill.
pSETΩ	Spectinomycin	pSET152 derivative in which the apramycin resistance cassette (*aac(3)IV*) is replaced with the spectinomycin resistance gene (*aad*) as a *SacI* fragment	(O'Connor, et al. 2002); Gift from J. Nodwell, McMaster University
pIJ6902	Apramycin	pSET152 derivative that contains the thiostrepton-inducible promoter *ptipA* from *Streptomyces*	Gift from M. Buttner, John Innes Centre
pIJ6902Ω	Spectinomycin	pIJ6902 derivative in which the apramycin resistance cassette (*aac(3)IV*) is replaced with the spectinomycin resistance gene (*aad*) as a *SacI* fragment	Gift from D. Bignell, University of Alberta
pAU227	Ampicillin	Entire helicase gene from *SacI* site 366 nt upstream of ATG start codon to *NcoI* site 352 nt downstream of stop codon. *SacI/NcoI* fragment was cloned directly into the same sites in pUC120 polylinker.	(Stoehr 2001)

Table 2.3 continued: Plasmids and cloning vectors used in this study

Plasmid	Antibiotic Resistance Marker	Relevant Characterisitcs	Reference (or Source)
***E. coli-Streptomyces* shuttle vectors**			
pAU316	Apramycin	*bldG* is present as an *NdeI-XbaI* fragment cloned into the similarly digested pIJ6902 vector; expression of *bldG* is under control of the p*tipA* promoter.	(Bignell, et al. 2003)

Table 2.4: Recombinant plasmids and cosmids constructed and used in this study

Plasmid/ Cosmid	Parent Plasmid/Cosmid	Selective Marker(s)	Insert size	Relevant Characteristics	Use
pAU330	pMAL-c2X	Ampicillin	2769 bp	Vector contains the entire *sco3550* coding region from the predicted ATG start codon to an *NcoI* site 352 nt downstream of the stop codon. Cloned in a two-part strategy (see Section 3.2.).	Overexpression of MBP-SCO3550 in *E. coli*
pAU331	pIJ6902	Apramycin	2769 bp	Vector contains the entire *sco3550* coding region from the predicted ATG start codon to an *NcoI* site 352 nt downstream of the stop codon. Cloned in a two-part strategy similar to pAU330 but 5' PCR product was cloned using primer JST12-2 and an engineered *NdeI* site and product was cloned into the *Nde* I/*Bam* HI sites of pIJ6902	Inducible overexpression of SCO3550 in *S. coelicolor*
pAU332	pAU331	Spectinomycin		pAU331 derivative in which the apramycin resistance cassette is replaced with a *SacI* digested spectinomycin resistance fragment	Complementation of putative *sco3550* deletion mutants. Positive control for western analysis, and conditional mutagenesis

Plasmid/ Cosmid	Parent Plasmid/Cosmid	Selective Marker(s)	Insert size	Relevant Characteristics	Use
pAU333	Cosmid SCH5	Ampicillin, Apramycin, Kanamycin	39.59 kb	Same as SCH5, but 2330 bp internal region of *sco3550* has been replaced with a 1384 bp fragment containing the (*aac(3)IV*) apramycin resistance cassette and the *oriT* of plasmid RP4	PCR-targetting deletion of *sco3550* in *S. coelicolor*
pAU334	pSETΩ	Spectinomycin	3200 bp	Vector contains entire *sco3550* coding region from pAU227 *SacI* site 366 nt upstream of ATG start codon to *NcoI* site 352 nt downstream of the stop codon. Insert overhanging ends were Klenow blunted and cloned into pSETΩ that had been digested with *EcoRI* and blunted.	Complementation of putative *sco3550* deletion mutants.
pAU335	pIJ6902Ω	Spectinomycin	473 bp	Vector containing *bldG* as an *NdeI/XbaI* fragment; expression of *bldG* is under contol of p*tipA* promoter.	Conditional mutagenesis of *sco3550.*
pAU336	pAU331	Apramycin	60 bp	pAU331 derivative where a 6x histidine encoded orf was cloned on the 5' end of *sco3550* as a *Nde* I fragment	Overexpression of recombinant SCO3550 in *S. coelicolor* and chemical cross-linking studies.

2.1.4 Preparation of *E. coli* glycerol stocks

E. coli strains were grown as an overnight liquid culture containing appropriate antibiotic selection and then mixed with an equal volume of 40% v/v glycerol. The stocks were flash frozen in a dry ice ethanol bath and stored at -85°C.

2.1.5 Growth and maintenance of *S. coelicolor* strains

Liquid cultures of *S. coelicolor* strains were grown in R2YE (Hopwood et al., 1985) or 2 × YT broth (Sambrook et al., 1989) as 5 mL cultures in universal bottles containing a 1 inch spring coil, or in R2YE broth as 25-50 mL cultures in 250 mL flasks containing a spring coil, and were incubated at 30°C while shaking at 200 rpm. Surface cultures of *S. coelicolor* strains were grown on either R2YE agar, Difco nutrient agar (DNA, 23 g/L) or MS agar (20 g/L soya flour, 20 g/L mannitol, 20 g/L agar) (Sambrook et al., 1989) and were incubated at 30°C. When necessary, the medium was amended with various antibiotics including, apramycin (25-50 μg/mL), nalidixic acid (25 μg/mL), thiostrepton (30 μg/mL) or spectinomycin (500-1000 μg/mL). Strains were maintained as solid plate cultures at 4°C or as either spore or mycelial stocks in 20% (v/v) at -20°C.

2.1.6 Preparation of *S. coelicolor* glycerol stocks

S. coelicolor spore stocks were prepared as described in Hopwood *et al.* (1985). For mycelial stocks, the cultures were grown on cellophane disks on the surface of R2YE agar for 3-4 days. The mycelia were then scraped from the cellophane disks into ground glass homogenizers containing sterile milli-Q water, and

the resulting suspension was homogenized. The homogenized supension was transferred to a universal bottle, centrifuged in an International Centrifuge with a swinging bucket rotor at 3000 rpm for 10 minutes at 4°C, the supernatant decanted and the mycelial suspension was added to an equal volume of 40% glycerol to yield a final concentration of 20%, aliquoted and stored at -20°C.

2.2 DNA Isolation and Transformation

2.2.1 Preparation of *E. coli* competent cells

Competent cells were prepared by inoculating LB liquid medium containing antibiotic, when necessary, and then growing the culture overnight at 30 or 37°C on a rotating rack. The culture was then diluted 1:100 with fresh LB medium and antibiotic, if necessary, and grown at 30 or 37°C while shaking at 200 rpm until an OD_{600} of approximately 0.6 was reached. For PCR-targeted mutagenesis of *sco3550*, electrocompetent cells of *E.coli* BW25113/pIJ790 containing Cosmid SCH5 were prepared after growth of the culture to an OD_{600} of approximately 0.6 in the presence of 10 mM L-arabinose in order to induce the expression of the λRed proteins. Culture was then centrifuged at 3000 rpm for 10 minutes at 4°C in an International Centrifuge with a swinging bucket rotor in order to pellet and subsequently wash the cells. For chemically competent cells, the cell pellet was washed once with 5 mL of chilled 100 mM $MgCl_2$, once with 5 mL of chilled 100 mM $CaCl_2$, and was resuspended in 100 mM $CaCl_2$ + 20% glycerol to a final volume of 1mL. The cell suspension was then aliquoted into 50-200 µL volumes, flash frozen in a dry ice ethanol bath and stored at -85°C. For preparation of electrocompetent cells, the cell

pellet was washed twice with chilled 10% glycerol and reuspended in 10% glycerol to a final volume of 50 μL-1 mL depending upon the DNA species being transformed. For PCR-targeted mutagenesis, a higher concentration of electrocompetent cells per electroporation was commonly used, therefore, the 50-100 μL total volume of resuspended electrocompetent cells was used immediately for transformation. However, for standard electroporation of ccc DNA, the washed cell pellet was resuspended in 1 mL of 10% glycerol and aliquots of 50-100 μL were then used immediately for electroporation or flash frozen in a dry ice ethanol bath and stored at -85°C.

2.2.2 Transformation of *E. coli*

For transformation of chemically competent *E. coli* cells, cell aliquots were thawed on ice and plasmid DNA (1-7.5 μL) was added to the suspension and gently mixed. The cell suspension was then allowed to sit on ice for 30 minutes, followed by a brief 90 second heat shock at 42°C and then returned to ice for 2 minutes. LB liquid medium was added to the cell suspension to a final volume of 1 mL and the culture was incubated on a rotating rack at 37°C for 1 hour. For electroporation of electrocompetent cells, the cells were thawed on ice and plasmid DNA was added to the suspension and gently mixed. The suspension was transferred to a chilled 0.2 cm electroporation cell and a Bio-Rad GenePulser II was used for the electroporation at settings of 200 Ω, 25 μF and 2.5 kV. Chilled LB liquid medium was immediately added to a final volume of 1 mL and the cell suspension was incubated on a rotating rack at 37°C for 1 hour. Upon completion of the 1 hour incubation, transformed

chemically or electrocompetent cells were plated in 100 µL aliquots on LB agar medium containing the appropriate antibiotic and 100 mM IPTG and 40 µg/mL X-gal when the transformed construct supported blue-white selection screening. Plate cultures were then incubated overnight at 30 or 37°C, where indicated.

2.2.3 Isolation of plasmid DNA from *E. coli*

Plasmid DNA was isolated from *E. coli* utilizing the alkali lysis method detailed by Sambrook *et al* (1989).

2.2.4 Preparation of *S. coelicolor* protoplasts

Protoplasts were prepared essentially as described by Hopwood *et al.* (1985), except that cultures were routinely grown in TSB or R2YE liquid medium supplemented with 2.5 M $MgCl_2$ and 20% glycine.

2.2.5 Transformation of *S. coelicolor*

Transformation of *S. coelicolor* protoplasts was performed according to instructions described by Hopwood *et al.* (1985).

2.2.6 *E. coli* conjugation into *S. coelicolor*

The *E.coli* strain ET12567/pUZ8002 harboring the *oriT*-containing plasmid to be transferred was prepared for conjugation by inoculating 2 mL LB liquid medium containing the appropriate antibiotic and growing the culture overnight on a rotating rack at 37°C. The culture was then diluted 1:100 in fresh LB liquid medium and

grown on a rotating rack at 37°C until an OD_{600} of 0.4-0.6 was reached. The cell suspension was then transferred to a sterile universal bottle and centrifuged in an International Centrifuge in a swinging bucket rotor at 3000 rpm and 4°C for 10 minutes. The cell pellet was washed twice with fresh LB liquid medium to remove any remaining extracellular antibiotic and then resuspended in 1 mL of fresh LB liquid medium. For conjugation into *S. coelicolor* spores, approximately 10^8 spores were suspended in 500 µL 2 × YT broth, heat shocked for 10 minutes at 50°C and then added to 500 µL of the washed *E. coli* cells in a 15 mL round bottomed plastic tube (Corning). For conjugation into mycelial fragments, mycelia from a 3-4 day old R2YE plate culture grown on a cellophane disk, were scraped into a glass homogenizer and homogenized in 2-3 mL of 20% glycerol. An aliquot of 500 µL of the mycelial fragments was then added to 500 µL of the washed *E. coli* cells. Both the spore-*E. coli* and mycelial-*E.coli* mixtures were centrifuged in a clinical centrifuge at 3000 rpm for 1 minute, the supernatant was decanted and the pellet was resupended in any residual liquid. A dilution series of the mixture was then plated on MS agar plates containing 10 mM $MgCl_2$ and the plates were incubated at 30°C for 16-24 hours. The plate culture was then overlayed with 1 mL of sterile milli-Q water containing 0.5 mg nalidixic acid, to inhibit *E. coli* growth, and the appropriate antibiotic for selection of the transferred plasmid and the plates were further incubated for 3-5 days at 30°C. Single exconjugants were then re-streaked onto MS agar containing nalidixic acid and the appropriate antibiotic followed by another round of growth on R2YE agar medium containing nalidixic acid and the suitable antibiotic (Sambrook et al,, 1989).

2.2.6 Isolation of chromosomal DNA from *S. coelicolor*

Chromosomal DNA isolation was performed using the Total DNA isolation Procedure 3 from Hopwood *et al.* (1985). Deviation from the procedure occurred in step 6, where 6 mL of phenol/chloroform was used for subsequent extractions. Also, the procedure was scaled down into 1.5 mL microfuge tubes at the first precipitation stage and the DNA-RNA pellet was redissolved in 500 µL of TE containing 40 µg/mL RNase. The remaining procedure was then scaled down appropriately for a small-scale isolation.

2.3 DNA Purification and Analysis

2.3.1 Restriction digestion and cloning of DNA

Restriction endonuclease digestion of plasmid DNA, chromosomal DNA and PCR products was performed as per the manufacturer's instructions (Roche or New England Biolabs) and as detailed by Sambrook *et al.* (1989). Klenow enzyme (Roche) was primarily used in blunting reactions as described by Roche. The reaction mixture consisted of approximately 1 µg of the target DNA, 1 mM dNTPs, 1x Blunting Buffer (50 mM Tris-HCl, pH= 7.5, 10 mM $MgCl_2$, 1 mM DTT, 50 µg/mL BSA) and 1 U of Klenow enzyme for 15 minutes at 37°C, followed by a 10 minute inactivation step at 65 °C. Ligation reactions were carried out with a standard insert to vector ratio of 3:1 for overhanging-end ligations, while blunt- overhanging-end and blunt-blunt ligations used a slightly higher insert to vector ratio. Ligation

reaction mixtures consisted of appropriate amounts of digested insert and vector DNA, 1 mM ATP, 1x ligation buffer (50 mM Tris-HCl, pH= 7.6, 10 mM $MgCl_2$, 5% PEG 8000, 1 mM DTT) and 0.5-1 U of T4 DNA ligase (Roche) to a final volume of 5-15 µL. For overhanging-end ligations, the reaction mixture was incubated at 15°C overnight, while blunt-end ligations were incubated at RT for 4 hours followed by a 15°C incubation overnight. Typically, half of the ligation mixture was transformed into competent *E. coli* cells while the remaining volume was stored at 4°C.

2.3.2 Polymerase Chain Reaction (PCR)

All oligonucleotide primers used for PCR are listed in Table 2.5. Polymerase chain reaction was used to amplify DNA fragments to be used in cloning reactions as well as to prepare probes for Southern hybridization analysis and S1 nuclease mapping. PCR was also used to screen transformants for the presence of insert DNA and to test for vector integration into *S. coelicolor* chromosomal DNA. Finally, PCR was employed to amplify cDNA for reverse transcription polymerase chain reaction (RT-PCR) analysis. Reactions were performed in 20-100 µL volumes in 0.6 mL tubes (Rose) using the Techne PHC-2 Thermocycler, or in 0.2 mL thin-walled tubes (Rose) using the Perkin Elmer GeneAmp PCR System 2400. PCR reactions utilized Taq polymerase (gift from M.A. Pickard, University of Alberta), EXPAND™ Long Template or High Fidelity enzymes depending upon the nature of the reaction. A standard PCR mixture consisted of 1-10 ng of plasmid DNA, 1 mM $MgCl_2$, 0.2 mM dNTPs, 1x Reaction buffer (Buffer 3 for Long Template enzyme, Buffer 3 for High Fidelity enzyme, 10× Taq buffer for Taq polymerase), 40 pmol of forward and

reverse primers (see Table 2.5), 0-6% DMSO and 2.5 U of enzyme. For PCR reactions involving the amplification of chromosomal DNA, 1μg of DNA was first denatured in 0.4 N NaOH and 0.4 mM EDTA in a final volume of 100 μL and incubated at RT for 10 minutes. The reaction was then precipitated on ice by the addition of 1/10 volume 3 M sodium acetate and 2 volumes of 95% ethanol, and the resulting pellet was resuspended in 10 μL sterile water. Sterile mineral oil was used to overlay reactions when the Techne PHC-2 Thermocycler was used, while a 104°C heated lid was used with the Perkin Elmer Thermocycler to avoid evaporation of the reaction mixtures. A typical PCR program contained a 5 minute denaturation stage at 95°C, with subsequent 30 cycles of 95°C denaturation for 30 seconds, a 30 second primer annealing step at approximately 5°C below the primer Tm, see 2.3.11) and a 1 minute extension stage at 68°C for either EXPAND™ Long Template or High Fidelity enzyme or 72°C for Taq polymerase. Upon completion of the 30 intervening cycles, a final 10 minute 68°C or 72°C extension step was used to assure complete extension of the amplified DNA fragments.

In order to construct the apramycin resistance disruption cassette with 39 nucleotide extensions homologous to regions downstream and upstream of *sco3550*, PCR was performed utilizing a gel-purified 1383 bp *Eco*RI/*Hind*III fragment from pIJ773 (see Table 2.3 and Section 2.3.4) as template and primers KGI2 and KGI3 (see Table 2.5). Reaction mixtures consisted of 50 ng template DNA, 50μM dNTPs, 50 pmol of each primer, 5% DMSO (warmed at 37 °C), 1x Reaction Buffer 2 (Roche) and EXPAND™ Hi-Fidelity polymerase (2.5U) to a final volume of 50 μL. The reaction included a two-step program where an initial 2 minute 95 °C denaturation

Table 2.5: Oligonucleotide primers used in this study

Primer	Sequence (5'-3')*	Region of Homology	Use
BKL80	cgcgctgcaggctcacgcagcttgggc	upstream of *sco3550* nt -323- -306 (*PstI* site)**	PCR amplification, *sco3550*- specific probe
BKL82	cgcgtctagacgacgtacttcagtgcc	*sco3550* nt +556- +539 (*XbaI* site)	PCR amplification for pAU330 construction
BKL87	gtgccggtggcgacgac	*sco3550* nt +207- +233	PCR for pAU336 construction internal S1, sequencing
BKL88	atgctcctggaccggctc	*sco3550* nt +1- +18	PCR amplification for pAU330 construction
BKL91	cggcaagcttgacacctcggcttgaag	upstream of *sco3550* nt -187- -170 (*HindIII* site)	PCR for *sco3550* 5' S1 mapping probe
hel-L	agaactttcacaacctctcggaa	*sco3550* nt +369- +391	PCR verification of mutagenesis
hel-R	gcatgtccggattggtcag	*sco3550* nt +490- +472	PCR verification of mutagenesis
JST5	cgccttggtcccattgtccgtca	upstream of *sco3550* -109- -87	PCR verification of mutagenesis
JST12-2	ggccatatgctcctggaccggctcgc	*sco3550* nt +19- 0 (includes start codon)	PCR for pAU331 construction, internal S1 probe
KGI2	tccccggcggagcccggctcccggccggaacc gggcatg*atccggggatccgtcgacc*	*sco3550* nt -36- +3 (includes start codon) plus flanking sequence of *aac(3)IV+oriT* cassette	PCR targetted mutagenesis
KGI3	ccgcaccacccctgccctctgcgcggtccctcc cgcct*atgtaggctggagctgcttc*	*sco3550* nt +2366- +2328 plus flanking sequence of *aac(3)IV+oriT* cassette	PCR targetted mutagenesis
KGI4	gacggagtcctccaggtcca	*sco3550* nt +1645- +1664	PCR verification of mutagenesis
KGI5	cgaactcggcctcgacatct	*sco3550* nt +1088- +1114	PCR verification of mutagenesis

Primer	Sequence (5'-3')*	Region of Homology	Use
KGI6	ccgggcttgcaggcgagtgag	internal sequence *aac(3)IV+oriT* cassette	PCR verification of mutagenesis
KGI7	gggatcgatcgggtcgg	upstream of *sco3550* nt -39- -23	PCR for 5' S1 mapping probe
KGI10	ggcttgcacctcacgtca	*sco3550* nt +19- 0 (includes start codon)	PCR for internal S1 mapping probe
KGI11	acaaggatccgctgcacaag	*sco3550* nt +2241- +2255	PCR for 3' S1 mapping probe
KGI12	tccttagtcatgttggtgcgctgcgtgctc	downstream *sco3550* nt +2485- +2467	PCR for 3' S1 mapping probe
LBI	gcgaattccatatgggcagcagccatca	5' HIS_6 fragment	PCR amplification for pAU336 construction, sequencing

**: Contains an engineered restriction endonuclease site
Primer sequence: Normal text- sequence complementary to *sco3550* locus
italic text- sequence complementary to *aac(3)IV+oriT* cassette

step was followd by 10 cycles of 95 °C denaturation for 45 seconds, 50 °C annealing for 45 seconds, and a 72 °C extension for 1.5 minutes. The second stage of the program consisted of an additional 15 cycles where the annealing temperature was increased to 55 °C and the program was concluded with a 72 °C 5 minute extension period. A portion of the reaction (5μL) was analyzed by agarose gel electrophoresis (see Section 2.3.3) and the remaining sample was then purified as per the manufacturer's instructions with a PCR purification kit (Qiagen). DNA was eluted in 12 μL of 10 mM Tris-HCl pH=8.5, a fraction was analyzed by agarose gel electrophoresis to examine recovery efficiency and the remaining PCR product was used in subsequent attempts at PCR targetting of the *S. coelicolor* cosmid.

2.3.3 Agarose gel electrophoresis of DNA

DNA fragments from 1-10 kb in size were resolved by electrophoresis on 1.0% agarose, 1× TBE (90 mM Tris, 89 mM boric acid, 2.5 mM Na_2EDTA) gels. Smaller DNA fragments ranging in size from 0.5-1 kb were separated by electrophoresis using 2.0% agarose, 1× TBE gels. *Pst*I-digested Lambda DNA (Roche) was used as a molecular weight marker for reference to DNA samples greater that 1 kb, whereas, Molecular Weight Marker V (Roche) was used for smaller DNA fragments. Prior to electrophoresis, samples were resuspended in 1/5 volume of loading dye containing 0.25% bromophenol blue and 40% sucrose. Electrophoresis typically occurred at approximately 108 V. Gels were stained with ethidium bromide and the DNA was visualized via UV light illumination.

2.3.4 DNA purification from agarose gels

DNA fragments larger than 1 kb were purified from 1.0% agarose, 1× TAE (40 mM Tris-acetate, 1 mM EDTA) gels using the trough purification method described in Zhen and Swank (Zhen and Swank, 1993). DNA fragments that were smaller than 1 kb were sometimes agarose gel-purified using the Qiagen gel extraction kit as listed in the manufacturer's instructions. DNA purified in this manner was primarily used as templates for PCR and PCR-targeted mutagenesis.

2.3.5 Polyacrylamide gel electrophoresis of DNA

DNA fragments between 0.08 and 0.7 kb in size were resolved by electrophoresis at 200 V on 5% polyacrylamide (29:1 acrylamide:N,N'-methylene bisacrylamide, BioRad), 1× TBE gel using a BioRad Protean II xi cell™ vertical gel system, or a Mini-Protean™ 2 or 3 gel system. Molecular Weight Marker V was used as a reference marker and samples were prepared and DNA fragments visualized as described in section 2.3.3.

2.3.6 DNA purification from polyacrylamide gels

DNA fragments smaller than 1 kb were generally purified from 5% polyacrylamide gels using the Crush and Soak procedure detailed by Sambrook *et al.* (1989). The purified sample was ethanol precipitated and redissolved in 5-15 μL sterile water.

2.3.7 DNA Sequencing

Manual DNA sequencing was performed using a modified chain termination method (Sanger et al., 1977) for use with the Thermo Sequenase Radiolabeled Terminator Cycle Sequencing Kit (USB). Reaction mixtures were established as per the manufacturer's instructions and were performed in either 0.6 mL tubes when using the Techne PHC-2 Thermocycler, or in 0.2 mL thin-walled tubes when the Perkin Elmer GeneAmp PCR System 2400 Thermocycler was used. Labelled fragments were separated on a 6% polyacrylamide sequencing gel (19:1 acrylamide:N,N'-methylenbisacrylamide (BioRad), 8.3 M urea, 1xTBE). Bands were visualized by overnight exposure to a Molecular Dynamics storage phosphor screen and the terminated radiolabeled fragments were visualized by a phosphorImager (Molecular Dynamics Model 445 SI) scan and analyzed with Imagequant™ software. Automated sequencing was performed utilizing an Amersham DYEnamic ET kit purchased from the Molecular Biology Services Unit (MBSU). The sequencing mix consisted of the template DNA of interest, either 100-500 ng of double-stranded plasmid, 600-1000 ng of restriction enzyme digested cosmid, or 1 ng/bp of PCR product, as well as 1 μL primer (5 pmol/ μL), 1 μL DMSO, 8 μL ET mix and sterile water to a final volume of 20 μL. Reactions proceeded within 0.2 mL thin-walled tubes in the Perkin Elmer GeneAmp PCR System 2400 with a standard reaction program consisted of a 5 minute 95°C denaturation stage followed by 25 cycles of a 30 second 95 °C denaturation step, a 15 second 50 °C annealing step and a 60 °C extension step. A high GC protocol was also used for primers that had a Tm of higher than 50 °C. This program consisted of a 5 minute 95 °C denaturation stage

followed by 25 cycles of a 30 second 95 °C denaturation step and a 2 minute 60 °C stage for both annealing and polymerase extension. Samples were prepared for electrophoresis by stopping the reaction with addition of 2 μL NaOAc/EDTA and precipitation on ice by the addition of 80 μL 95% ethanol. Samples were only allowed to precipitate for 15 minutes in order to avoid the precipitation of free nucleotide which adds background fluorescence in the detection system. Samples were pelleted, washed with 500 μL of 70 % ethanol, stored at – 20 °C and provided to MBSU to be separated by electrophoresis

2.3.8 Southern hybridization analysis

Southern analysis (Southern, 1975) was performed according to the procedure of Hopwood *et al.* (1985). The DNA was cross-linked to the membrane on Program C3 (150 mJoules) using a Bio-Rad GS Gene Linker and the membrane was stored at either room temperature or – 20 °C.

2.3.9 Colony hybridization analysis

Colony DNA hybridization was utilized for large-scale screening of *E. coli* for identification of correctly cloned insert DNA. Circular nylon membranes (Hybond™ -N, Amersham) were labeled with numbered grids and placed on LB agar medium containing the appropriate antibiotics. Transformants of interest were streaked on the membrane and re-streaked upon LB agar medium without a nylon membrane and including an acetate numbered grid attached to the back of the plate. Plate cultures were incubated overnight at 37 °C. Membranes were then removed from the plate

and placed, colony side up, onto Whatman paper saturated with denaturing solution (1.5 M NaCl, 0.5 M NaOH) for 5 minutes. Membranes were neutralized by transfer onto Whatman paper soaked in neutralizing solution (3 M NaCl, 0.5 M Tris-HCl) for 2 segments of 5 minutes. Neutralized membranes were subsequently placed in a reservoir of 2 x SSC solution and the remaining cellular debris was scrubbed away, blotted dry on Whatman paper and DNA was cross-linked to the membrane on Program C3 (150 mJoules) using a Bio-Rad GS Gene Linker. Membranes were stored at either room temperature or – 20 °C.

2.3.10 Labeling of DNA probes

DNA fragments used as probes for Southern or Colony hybridizations (see 2.3.8-2.3.9) were labeled using the random primer labeling procedure described by Feinberg and Vogelstein (1983) and modified by Roche (Feinberg and Vogelstein, 1983). The DNA fragment (2 µL), which acted as a template for the internal labeling of fragments, was brought to a final volume of 9 µL in a 1.5 mL microfuge tube and denatured at 95 °C for 5 minutes. Reactions were incubated at 37 °C for 4 hours or at room temperature overnight. Unincorporated nucleotides were removed from the labeling reaction, in order to minimize background radioactivity, using a Micro Biospin® 6 chromatography column (Bio-Rad) as described by the manufacturer. To determine the correct volume to be added to hybridizations, 1 µL of labeled probe was Cerenkov counted in a Beckman LS 3801 scintillation counter.

DNA oligonucleotides or fragments used as radioactive primers for RT-PCR (see 2.4.4) and primer extension (see 2.4.3) or probes for 5' S1 nuclease mapping (see

2.4.2) were labeled at their 5' ends using the end-labeling procedure detailed by Chaconas and van de Sande (1980) modified by Roche (Chaconas and van de Sande, 1980). End-labeling reactions proceeded in a reaction mixture containing 1-3 μL of DNA (10-50 pmol oligonucleotide primer), 1× kinase buffer (50 mM Tris-HCl, pH=8.0, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine) and 5 μL (50 μCi) γ-^{32}P-ATP (4500 Ci/mmol, ICN or MP biochemicals). T4 polynucleotide kinase (10 U/μL, Roche) was diluted 1/10 in kinase dilution buffer (50 mM Tris-HCl, pH=7.6, 1 mM DTT, 0.1 mM EDTA, 50 % glycerol), and 1 μL was added to the reaction mixture and incubated for 15 minutes at 37 °C, followed by a second 1 μL addition and a further 15 minute incubation. Upon completion of the reaction, the reaction volume was increased to 50 μL and was precipitated on ice by addition of 1/10 volume 3 M sodium acetate, 2 volumes of 95 % ethanol and 2 μL glycogen. Probe preparation for 5' S1 nuclease mapping consisted of Cerenkov counting the entire DNA pellet and redissolving in sterile milli-Q water to final activity of 50 000 cpm/μL. End-labeled probes used in primer extension were redissolved in sterile milli-Q water to a final concentration of 2.5 pmol/μL.

DNA fragments used as probes for 3' S1 nuclease mapping were labeled at their 3' ends using a modified Roche procedure for random primer labeling. DNA (10-20 μL) was digested with an appropriate restriction enzyme to leave 3' recessed ends. The labeling reaction mixture consisted of the addition of 1 μL of 5 mM dNTPs (dATP, dGTP, dTTP), 2.5 μL (25 μCi) of α-^{32}P-dCTP (Amersham) and 1 U of Klenow enzyme (Roche) and the reactions were incubated at room temperature for

30 minutes. The labeled probe was then prepared as described for 5' S1 nuclease mapping.

2.3.11 DNA hybridizations

Southern and Colony DNA hybridizations were typically performed in glass hybridization bottles (Robbins or Fischer Scientific) and pyrex dishes, respectively. Hybridization buffer consisted of 3 X SSC, 4 X Denhardt's solution (0.08 % polyvinylpyrrolidone, MW 360 000), and 100 µg/mL denatured salmon sperm DNA in a final volume of 10 mL. An appropriate percentage of deionized formamide was included in the hybridization buffer to lower the effective hybridization temperature when the melting temperature of the probe was greater than 65 °C. Nylon membranes were pre-hybridized in a hybridization oven (Robbins or Fischer Scientific) for at least 1 hour prior to the addition of radiolabeled DNA probe (2 000 000-4 000 000 cpm/10mL) that had been denatured at 95 °C for 5 minutes. The hybridization was performed at approximately 25 °C below the melting temperature of the probe as defined by the formula, Tm= 81.5 °C + 16.6log$M$ + 0.41 (%G+C)- 500/n – 0.61 (% formamide), where M is the ionic strength (0.45 for 3 x SSC) and n is the shortest length of homology (Hopwood et al., 1985). For oligonucleotide probes, the hybridization temperature was approximately 5°C below the probe melting temperature as defined by the formula, Tm= 4(G+C) + 2(A+T) where G,C,A,T represent nucleotide composition of the probe (Hopwood et al., 1985). Hybridizations were incubated from 8-24 hours until the hybrization buffer was removed and the membranes were first washed by agitation with wash solution I (2 x

SSC, 0.1 % SDS), and then washed at the hybrization temperature with wash solution I twice for 20-30 minutes each and then washed twice with wash solution II (0.2 x SSC, 0.1 % SDS) twice for 15-30 minutes at 55-65°C. Membranes were blotted dry on Whatman chromatography paper, wrapped in Saran wrap, exposed to a phosphor screen and visualized with a phosphorimager.

Hybridized probes were removed from Southern membranes by placing the membrane in boiling 0.1 % SDS and allowing the solution to cool to room temperature while gently shaking. Membranes were stripped with at least 2 washes of 0.1 % SDS until the radioactive signal was sufficiently reduced and the stripped membrane was then blotted dry, covered with Saran wrap, exposed on a phosphor screen and visualized with a phosphorimager in order to determine the efficacy of stripping. Completely stripped membranes were then stored at – 20 °C.

2.4 RNA Analysis

2.4.1 Isolation of RNA

RNA was isolated from *S. coelicolor* using a modified version of Kirby *et al.* (1967) (Kirby et al., 1967) procedure as described by Hopwood *et al* (1985). RNA isolated from plate cultures was provided as gifts from either D. Bignell, C. Galibois or K. Colvin. RNA isolated from liquid cultures entailed growing the *Streptomyces* in initial seed culture containing 5 mL of R2YE broth in a sterile spring containing universal (Section 2.1.5). Seed cultures were grown overnight at 30 °C while shaking at approximately 240 rpm. The entire inoculum from the seed cultures was used to inoculate 25-50 mL R2YE broth cultures in 250 mL spring Erlenmeyer flasks and the

culture was grown to a high density of growth at 30 °C while shaking at 240 rpm. Processing of the samples began with pelleting the mycelia by centrifugation in an International Centrifuge with a swinging bucket rotor at 3000 rpm for about 30 seconds. Centrifugation was done at room temperature to avoid cooling the mycelia in order to prevent the expression of a subset of cold shock genes. The pelleted mycelia were then washed with DEPC treated water. After centrifugation and decanting of the wash, 5 mL 1× Kirby's Mix (1% sodium tri-isopropylnaphthalene sulfonate, 6% 4-amino salicylate, 6% neutral phenol buffered in 50 mM Tris-HCl pH=8.3) and glass beads (Fisher, 3 mm) were added directly to the mycelia and the mixture was vortexed for 4 × 30 seconds with intermittent cooling. Additional phenol/chloroform (5 mL) was added to the mixture and was vortexed for 1 minute and then the entire emulsion was transferred into 13 mL polypropylene tubes (Sarstedt). Phase separation occurred by centrifugation at 8500 rpm in a Beckman J2-H5 centrifuge with the JA20 rotor, and the upper aqueous phase was transferred to a new polypropylene tube. Extractions with 5 mL of phenol/chloroform were performed twice more and the remaining nucleic acid was then precipitated in a fresh polypropylene tube by the addition of 1/10 volume 3 M sodium acetate and 1 volume isopropanol at –80 °C overnight. Samples were thawed on ice and pelleted by centrifugation at 8000 rpm for 10 minutes in the Beckman centrifuge. Pellets were washed with 95 % ethanol and after careful drying resuspended in 450 µL of DEPC-treated sterile milli-Q water or UltraPure DNase-RNase free water and transferred to RNase-free 1.5 mL microfuge tubes (Fisher). Contaminating DNA was removed by a DNase treatment after the addition of 1/10 volume 10x DNase buffer (0.5 M Tris-HCl

pH=7.8, 0.05 M $MgCl_2$) and 2 separate aliquots of 7 μL RNase-free DNase (Roche, 10 U/μL) at 30 minute intervals. Upon completion of the 1 hour DNase incubation, remaining protein was removed from the solution by 2 separate phenol/chloroform extractions, followed by 3 chloroform extractions. RNA was precipitated by addition of 1/10 volume 3 M sodium acetate and 1 volume isopropanol and left on ice for 30 minutes. Pellets were briefly washed with 95 % ethanol and redissolved at 4 °C in up to 200 μL of DEPC-treated sterile milli-Q water. Initial quantification of RNA was defined by spectrophotometric analysis measuring the absorbance at 260 nm and quality of the RNA was analyzed by agarose gel electrophoresis. RNA samples were stored over long durations of time as a sodium acetate and isopropanol precipitate at –80 °C.

2.4.2 High resolution S1 nuclease mapping

Both 5' and 3' S1 nuclease mapping of *sco3550* transcripts was performed as previously described (Bignell et al., 2000). Probes were designed to discriminate between full-length protection of the probe and probe-probe reannealing; therefore, probes were modeled to span the predicted transcription start or stop sites and included non-homologous extensions at the 3' end, for 5' S1 nuclease mapping, or at the 5' end, for 3' S1 nuclease mapping. A *hrdB* internal control, specific to a constitutively expressed sigma factor in *S. coelicolor*, was also utilized to control for RNA integrity and RNA loading. Probes (100 ng) were end-labelled using the standard end-labeling procedure (see 2.3.10), precipitated and redissolved to a final concentration of 30 fmol/μL in RNase-free water. The amount of ng per fmol of the

probe was calculated by: N bp x 660 Da/bp x (1 mol/10^{15} fmol) x 10^{9} ng/g = ng/fmol).

For both 5' and 3' S1 mapping, 40 µg of total RNA was pelleted from the stored isopropanol suspensions and 1µL of the end-labelled probes (30 fmol/µL) were added to the RNA pellet and dried in a Speed Vac Concentrator (Savant) for approximately 2 minutes. Note that a probe only control (designated probe+S1), in the absence of total RNA, was also dried down at this stage. Immediately after drying, 20 µL of NaTCA hybridization buffer (3 M NaTCA, 50 mM PIPES, 5 mM EDTA pH=7.0) was added to the dried pellet and vortexed. Repeated rounds of 2 minutes heating at 70 °C, vortexing and pulsing in a microfuge were performed in order to ensure complete redissolving of the pellet. Once the pellet was completely dissolved the reactions were heated at 70 °C for 15 minutes to completely denature the DNA probe. After this brief denaturation period, the reactions were transferred directly from the temperature block into a controlled temperature incubator at 45°C and incubated overnight to allow probe-transcript, or probe-probe in the control, annealing. Because the difference between probe/transcript and probe/probe melting temperatures is large enough in NaTCA buffer, all hybridizations can be performed at 45 °C without experimentally determining the melting temperature of the probe. Digestion with S1 nuclease was performed by addition of 100 U of S1 nuclease (Roche), 60 µL 5x S1 digestion buffer (0.28 M NaCl, 30 mM CH_3COONa pH= 4.4, 4.5 mM $(CH_3COO)_2Zn$, 20 µg partially-cleaved denatured calf thymus DNA) and 240 µL RNase-free water to a final volume of 320.25 µL, followed by brief vortexing, pulsing and incubation at 37 °C for 45 minutes. Reactions were halted by the

addition of 75 µL S1 termination solution (2.5 M ammonium acetate, 0.05 M EDTA) and the nucleic acid duplex was precipitated on ice for at least 10 minutes by adding 50 µg glycogen (Roche) and 400 µL isopropanol. Nucleic acids were pelleted, washed with 75 % ethanol and fully redissolved in 3 µL of loading dye (see 2.3.5). Standard manual sequencing reactions were established (see 2.3.7) using the internal primer lacking the non-homologous end as a primer in the case of 5' S1 nuclease mapping or a primer just upstream of the 3' end of the probe in 3' S1 mapping. The S1 nuclease-digested samples and sequencing reactions were denatured at 95 °C for 5 minutes and separated by electrophoresis on a 6% polyacrylamide sequencing gel for up to 3.5 hours at 40 W constant power. The gel was then removed from the sequencing apparatus, dried under a vacuum and exposed to a phosphorscreen for visualization.

2.4.3 Primer extension analysis

For primer extension analysis, 40 µg of RNA was removed from the storage isopropanol precipitate, pelleted, washed with 80% ethanol and redisolved in 9.76 µL RNase-free water. Reactions followed the modified version of C. Therm. Polymerase One Step RT-PCR System by Roche and included in a 20 µL final volume (10.24 reaction mix + 9.76 µL redissolved RNA), 1x RT buffer (Roche), 5 mM DTT, 2 mM dNTPs, 3.2 % DMSO, 5 pmol of end-labelled RT primer, 17.5 U RNA Guard (Amersham) and 15 U of C. Therm. polymerase. Primer extension reactions were performed in 0.2 mL thin-walled tubes when using the Perkin Elmer GeneAmp PCR System 2400 Thermocycler, and typical reaction conditions consisted of a 1 hour

extension period at 59-69 °C depending upon the desired stringency of the reaction, followed by a 10 minute 80 °C period of denaturation of the radiolabelled DNA/RNA hybrid. The reaction mixture was transferred to 1.5 mL microfuge tubes and precipitated on ice for 30 minutes by the addition of 40 μg glycogen, 1/10 volume 3 M sodium acetate, 2 volumes 95 % ethanol to a final volume of 157 μL. Samples were centrifuged and pellets, after a brief ethanol wash, were redissolved in 3 μL loading dye (stop solution). The samples were then denatured at 95 °C for 5 minutes and at least half the reaction volume was loaded onto a 6 % sequencing gel along with sequencing reactions generated using the same primer used for the primer extension reaction. Samples were separated by electrophoresis for up to 2 hours at 40 W constant power, the gel was removed from the sequencing apparatus, dried and exposed to a phosphorscreen for visualization. Due to the discontinuation of C. Therm. Polymerase One Step RT-PCR System by Roche, later attempts at primer extension analysis utilized a two-step reverse transcription reaction carried out with the Superscript™ III (Invitrogen) enzyme as described (see below).

2.5 *E. coli* Protein Overexpression and Purification

2.5.1 Overexpression of MBP-SCO3550

SCO3550, the putative RNA helicase, was overexpressed in *E. coli* DH5α and BL21 (DE3) as an N-terminal maltose binding protein-fusion protein (MBP). Initially, a small-scale pilot experiment was performed as detailed in the pMAL-c2X instruction manual [Protocol A, NEB (New England Biolabs)] using both, *E. coli* DH5α and BL21 (DE3) harboring the IPTG inducible pAU330 (see Table 2.4), to

ascertain if the fusion protein was being expressed and to what degree in a soluble form. A large-scale expression of MBP-SCO3550 was performed in the protease deficient *E. coli* strain BL21 (DE3) harboring the IPTG inducible pAU330. A single colony of the aforementioned strain was used to inoculate 2 mL LB broth containing 0.2% w/v glucose and 100 μg/mL ampicillin. The culture was incubated on a rotating rack for approximately 8 hours at 37 °C and was then diluted 100 fold into 10 mL of fresh LB broth plus glucose and ampicillin and incubated overnight at 37 °C. The 10 mL overnight culture was then used to inoculate 500 mL- 1 L LB broth containing glucose and ampicillin that was grown at 30°C with shaking at 240 rpm until the OD_{600} was 0.5-0.6. IPTG was then added to the large-scale culture to a final concentration of 0.3 mM to induce fusion protein overexpression and the culture was grown at room temperature while shaking at 240 rpm for 6 hours. Note that the use of a protease deficient BL21 (DE3) strain, as well as the times and temperatures of growth and induction were chosen to maximize the soluble expression of the fusion protein. The cells were pelleted by centrifugation at 4000 × g for 20 minutes in a Sorvall centrifuge, resuspended in 50 mL column buffer (20mM Tris-HCl pH=7.4, 200 mM NaCl, 1 mM EDTA) containing 1× Complete EDTA-free protease inhibitor cocktail (Roche) and stored at –20 °C overnight.

2.5.2 Purification of MBP-SCO3550

Purification of the overexpressed fusion protein, MBP-SCO3550, was undertaken using a slightly modified version of amylose resin affinity chromatography as described by the NEB pMal-c2X instruction manual. Cells that

were resuspended in 50 mL column buffer and stored overnight at –20 °C were thawed in cold water and the cell suspension was lysed by sonication for 2-4 × 30 seconds using a 2.5 mm probe on setting 1 (low) (Branson Sonifier 450). The resulting cell lysate was centrifuged at 9000× g for 30 min at 4°C and the supernatant, containing soluble fusion protein and other soluble species in the crude extract, was transferred to a sterile container and diluted 1:4 in column buffer. Large-scale purification was performed in a 60 mL syringe (Becton Dickinson and Co.) containing 15 mL of amylose resin, while smaller scale purification of fusion protein for enzyme assays was performed with smaller portions of cell extract and purified using 500 µL- 1 mL of amylose resin in disposable columns (Bio-Rad). Note that the amylose resin was stored at 4 °C in an equal volume of ethanol which had to be removed by washing the resin with at least 8 column volumes of column buffer prior to the addition of the diluted crude extract. Once the diluted crude extract was added to the column, the resin was washed with 12-25 column volumes of column buffer. Finally, soluble fusion protein was eluted from the amylose resin by the addition of column buffer containing 10 mM maltose and 0.1-0.25 column volume fractions were collected and analyzed by SDS-PAGE (see below). MBP-SCO3550-containing fractions composed of few other contaminating proteins, were pooled, quantified by Bradford Dye-binding assay (see below) and stored at –85 °C in 50% v/v glycerol.

2.5.3 Preparation of *S. coelicolor* cell-free extracts

Streptomyces plate culture cell-free extracts were prepared for Western analysis by growing the strain of interest on the surface of cellophane disks on R2YE

plates containing the appropriate antibiotic for varying lengths of time, most often corresponding to distinct developmental phenotypic transitions. Mycelia were scraped from the disks into 1.5 mL microfuge tubes with a sterile spatula and the biomass was suspended in a maximum of 1 mL lysing buffer [50 mM HEPES pH=7.2, 1× Complete EDTA-free protease inhibitor cocktail (Roche)]. The suspension was lysed by sonication on ice for a maximum of 5× 15 seconds at setting 1 using a 2.5 mm probe (Branson Sonifier 450). The cellular debris was separated by centrifugation at 13 200 rpm at 4 °C for 10 minutes and the supernatant, containing the soluble cell-free extract, was pooled, quantified (see below) and distributed into smaller aliquots and stored at –85 °C. *Streptomyces* liquid culture cell-free extracts were prepared for Western analysis and immunoprecipitation experiments similarly to those from plate cultures. In this case however, 25-50 mL R2YE liquid cultures were grown to a high cell density or for a specific amount of time, centrifuged at 3000 rpm for 10 minutes in an International Centrifuge with a swinging bucket rotor and washed once with sterile milli-Q water, before being resuspended in lysing buffer and further processed as described above.

2.5.4 Overexpression of SCO3550 in *S. coelicolor*

S. coelicolor strains were used that harbored recombinant plasmid pAU331, pAU332, or pAU336 (see table 2.4) that all contained *sco3550* under the control of thiostrepton inducible p*tipA*. The strains were grown to a high OD (prior to the onset of physiological differentiation) in 25-50 mL R2YE broth at 30 °C while shaking at 240 rpm and induced with thiostrepton (30 μg/mL final concentration). Standard

induction proceeded with a further 3 hour incubation at 30 °C while shaking at 240 rpm and cell-free extracts were prepared using the standard protocol (see Section 2.5.3).

2.5.5 Purification of HIS_6-SCO3550 from *S. coelicolor*

Due to the low level overexpression from p*tipA* relative to *E. coli* overexpression systems, a batch purification under native conditions technique (Novagen) was performed for the purification of recombinant SCO3550 from *S. coelicolor*. *S. coelicolor* strain M600 harboring pAU336 was grown and *sco3550* was induced as described in section 2.5.4. Ni-NTA resin (Qiagen) was prepared by washing 1 mL of the Ni-NTA/ ethanol slurry with at least 3 × 4 mL 1× Ni-NTA Bind Buffer (Novagen; 300 mM NaCl, 50 mM sodium phosphate buffer, 10 mM imidazole pH=8.0). Note that 1× Ni-NTA bind buffer contains 10 mM imidazole to minimize binding of contaminating proteins and thereby increase the purity of the final elution fractions. Washing was performed by brief hand mixing and separation by a 1 minute centrifugation at 1000 rpm in an International centrifuge with swinging bucket rotor. Subsequently, approximately 10 mg of *S. coelicolor* crude extract protein containing overexpressed recombinant HIS_6-SCO3550 was added to the washed and pelleted resin and mixed gently by shaking at 4 °C for 1- 3 hours. The lysate-Ni-NTA mixture was added to a 5 mL column (BioRad) with the bottom outlet capped and the resin was allowed to settle. Once settled, the bottom cap was removed and the column flow-through was collected for SDS-PAGE analysis. In order to remove non-specific interacting proteins, the column was washed with 2-4 × 4 mL 1× Ni-NTA wash

buffer (Novagen; 300 mM NaCl, 50 mM sodium phosphate buffer, 20 mM imidazole pH=8.0) and wash fractions were also collected for future analysis. Recombinant SCO3550 was eluted by addition of 4 x 0.5 mL 1x Ni-NTA elution buffer (Novagen; 300 mM NaCl, 250 mM imidazole, 50 mM sodium phosphate buffer pH=8.0) and elution fractions were collected.

2.5.6 Protein quantification

Quantification of proteins was performed using the procedure detailed by Bradford (1976) and modified by the manufacturer's instructions of the Bio-Rad Protein Assay Kit (Bradford, 1976). Bovine γ-Globulin was used as a standard.

2.5.7 Protein Precipitation

Protein samples were tri-chloroacetic acid (TCA) precipitated in order to concentrate dilute cell-free extract samples prior to SDS-PAGE or to equalize buffer discrepancies between purified protein samples and crude extract. Samples were precipitated on ice for 30 minutes in 0.25 volumes of 50 % TCA and the resulting precipitate was pelleted by centrifugation at 14 000 × g for 30 minutes at 4 °C. The pellet, which was not always entirely visible, was washed with ice cold acetone, pelleted to remove remaining acetone and air-dried. Pellets were redissolved in 2× SDS-loading dye (0.15 M Tris-HCl pH=6.8, 5 % SDS, 24 % glycerol, 12 % 2-mercaptoethanol, 0.0024 % bromophenol blue) to dissolve the pellet. To avoid acid hydrolysis of protein samples, especially during the 95 °C boiling stage, any

remaining acid was neutralized by the addition of 1 M Tris-HCl pH=8.0, until a color shift arose in the bromophenol blue pH indicator from yellow to blue.

2.6 Protein analysis

2.6.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

S. coelicolor cell-free extracts and purified fusion proteins were analyzed by SDS-PAGE as detailed by Sambrook *et al.* (1985) with some modifications. Protein samples were prepared by the addition of 1/4 volume of 2× SDS loading dye to a final volume of less than 35 µL when the Mini-Protean™ 2 (or 3) gel system was used or 200 µL with the BioRad Protean II xi cell™ vertical gel. Samples were denatured at 95 °C for 10 minutes prior to electrophoresis. SDS-PAGE broad range standards (1 µL, BioRad) or Kaleidoscope Pre-stained standards (10-15 µL, BioRad) were used as molecular weight markers for SDS-PAGE analysis. Samples were separated by electrophoresis on 6 or 7.5 % SDS-PAG in the case of SCO3550 analysis. The resolving gel contained 6-7.5 % polyacrylamide [(29:1 acrylamide: N,N' methylene bisacrylamide), 0.375 M Tris-HCl pH= 8.8, 0.1 % SDS)] and the stacking gel contained 3.2 % polyacrylamide [(29:1 acrylamide: N,N' methylene bisacrylamide), 0.125 M Tris-HCl pH= 6.8 , 0.068 % SDS]. Tris-glycine-SDS buffer (0.05 M Tris, 0.38 M glycine, 0.1 % SDS) acted as the protein running buffer and samples were separated by electrophoresis at room temperature at a constant voltage of 200 V for mini gels or a constant current of 40 mA for large gels. Upon completion of electrophoresis, gels were stained for 20- 30 minutes at 37°C in Coomassie Brilliant Blue protein stain [10 % glacial acetic acid, 50 % methanol, 0.001-0.002 % w/v

Coomassie Brilliant Blue (Bio-Rad)], rinsed with distilled water and destained in a solution of 25 % methanol and 10 % acetic acid either overnight at room temperature or for a couple hours at 37 °C. Gels with sufficient destaining were processed for long term storage by soaking in 5 % glycerol, followed by transfer to Fisherbrand thick chromatography paper, and drying for 1 hour under vacuum on a gel-dryer (Bio-Rad). Alternatively, if the bands of interest were bright enough, the gel was simply photographed on a light box with a digital camera.

2.6.2 Preparation of poly-clonal antibodies

Polyclonal antibodies raised against purified MBP-SCO3550 were prepared through successive rounds of injection of 250-500 μg of purified fusion protein with an equal volume of adjuvant. Fusion protein was purified through resolving by 7.5 % SDS-PAGE for 5-6 hours at a constant current 40 mA. Gels were then stained for 30 minutes at room temperature in 0.05 % Coomassie blue in water. Gels were subsequently washed in distilled water to remove the Coomassie blue and to partially destain the gel and the appropriate band were excised and transferred to 1.5 mL microfuge tubes and frozen at –20 °C. After a brief thawing period, the gel pieces were crushed and 1× PBS buffer was added to the tube up to 0.75 mL final volume. The emulsion was then passed through a series of 18.5, 21 and 23 gauge syringe needles in order to shear the gel pieces and the resulting suspension was combined with an equal volume of Freund's complete adjuvant for the initial injection, and Freund's incomplete adjuvant for successive rounds of injections. Two rabbits were injected at 4 week intervals to a total of 5 separate injections over a 16 week period

and blood samples were received 2 weeks after each injection in order to ascertain the specificity and titre of the serum using western analysis as described below. Serum was prepared from blood samples by an initial incubation for 1 hour at 37 °C to clot the blood followed by sedimentation of the clotted blood while standing overnight at 4 °C. Cleared serum was removed from the sedimented cellular debris with a long-stem pasteur pipette and remaining debris was removed by centrifugation of the serum sample in a Beckman centrifuge at 10 000 × g. The serum was aliquoted into smaller volumes and stored at –80 °C.

2.6.3 Western hybridization analysis

Immunodetection of SCO3550 protein from *S. coelicolor* cell-free extracts was performed by western analysis. Initially, 40-50 µg of *S. coelicolor* cell-free extract was separated by 6- 7.5 % SDS-PAGE as described in section 2.6.1. Kaleidoscope pre-stained standards (Bio-Rad) were used in order to serve as a molecular weight marker for migration on the polyacrylamide gel, as an indicator of efficacy of transfer to a PVDF (Bio-Rad) nylon membrane, and for subsequent size comparison with the final hybridized bands. Upon the completion of electrophoresis, the stacking gel was separated from the resolving gel and the resolving gel was soaked in transfer buffer (19.2 mM Tris-HCl pH=8, 0.192 M glycine, 0.015 % SDS, 20 % methanol) for 30 minutes at room temperature on an orbital shaker or rocker. During this incubation period a hydrophobic Sequi-Blot PVDF membrane (Bio-Rad) was briefly wetted in methanol and equilibrated in transfer buffer for at least 15 minutes. Transfer of proteins to the membrane was established in a small BioRad

Transblot apparatus at 100V (<400 mA for transfer of 2 gels, <350 mA for transfer of 1 gel) for 1 hour at 4 °C. After transfer, the membrane was initially washed 3× 5 minutes in sterile water, once for 5 minutes in Tris-buffered saline-tween (TBS-T) buffer (20 mM Tris-HCl pH= 7.6, 140 mM NaCl, 0.5 % Tween-20), and then blocked for at least 1 hour at room temperature in 20 mL blocking buffer (10% skim milk in TBS-T blocking buffer). The primary hybridization was set up to probe for SCO3550 by removal of 10 mL of blocking buffer (to be saved at 4 °C for the secondary hybridization) followed by the addition of 20 µL anti-sera, raised against SCO3550 (1:2000 dilution), to the remaining 10 mL blocking buffer and further incubation at room temperature for 1 hour. The membrane was briefly rinsed twice, washed 2 × 5 minutes and 1 × 15 minutes in TBS-T, and then incubated for at least 1 hour with horseradish peroxidase-conjugated secondary antibody (Amersham, 1:10 000) in the previously stored 10 mL blocking buffer. The membrane was washed with TBS-T buffer as previously described and bound secondary antibody was detected by incubating the blot in Western Lightning Plus (Perkin Elmer) reaction mixture of chemiluminescent reagents for 1 minute, followed by exposure of the membrane for 1-2 minutes until bands were clearly visible. Note that for anti-BldG western hybridization analysis, 15 µg of cell extract was separated by 15 % SDS-PAGE and probed with primary OG1 antibody at 1 in 20 000 dilution.

The membranes were periodically stripped and reused. Stripping involved an initial 5 minute wash in sterile milli-Q water, a single 5 minute wash in 0.2 N NaOH and a final 5 minute wash in water.

2.6.4 Affinity purification of poly-clonal antibodies

To enrich for antibodies specific to the MBP-SCO3550 fusion protein to remove nonspecific antibodies from the serum, affinity purfication was performed. Purified fusion protein (10 μg) and Kaleidoscope prestained standards were separated, in duplicate, by 7.5 % SDS-PAGE and transferred to a PVDF membrane using the standard electroblotting procedure (see 2.6.3). The membrane was split in half and one section was briefly stained in Ponceau S dye to correctly position the migration of the fusion protein relative to the prestained standards. Consequently, the correct location of the fusion protein on the unstained membrane was determined and approximately 1 cm square of the corresponding membrane fragment was excised and incubated in 10 % skim milk blocking buffer while shaking at room temperature for 1 hour. The membrane fragment was then transferred to a 1.5 mL microfuge tube, washed 3 × 5 minutes in 500 μL 1 × PBS buffer and incubated overnight with an equal volume of undiluted anti-serum and 1 × PBS buffer to a total volum of 500 μL. The membrane was then washed as previously described in 1 × PBS buffer and bound antibodies were eluted from the membrane through 2 × 30 minutes incubations with 200 μL 0.1 M glycine buffer pH= 2.0 while shaking and periodically mixing. Elution was followed by immediate neutralization with 30 μL 1 M Tris-HCl pH=8.0 and the titer and specificity of the affinity-purified antibody was then tested by western analysis of *S. coelicolor* cell-free extracts as descibed in section 2.6.3, typically using 1:200 – 1:2000 dilution of the affinity-purified antibody. BSA and glycerol were added to 0.1 mg/mL and 50 % final volume, respectively, and affinity purified antibodies were stored at – 80 °C.

2.7 *in vitro* Biochemical analysis

2.7.1 RNA helicase assays

RNA unwinding activity of SCO3550 was tested using *in vitro* RNA unwinding assays that were performed as described in Yu and Owttrim (2000), with partial double-stranded radioactive substrates prepared and provided by the Owttrim lab (Yu and Owttrim, 2000). Typically, reactions were performed for 30 minutes at 37 °C in a 20 µL reaction volume containing 20 mM HEPES-KOH pH=8.3, 3 mM $MgCl_2$, 1 mM DTT, 5 mM ATP, RNA guard, 500 µg/mL BSA, approximately 2000 cpm partially double-stranded RNA (approximately 50 fmol) and 600 ng recombinant SCO3550. Reactions were stopped by the addition of 5 µL 2× SDS-loading dye and reaction products were separated by 10 % SDS-PAGE. The gel was then dried on a BioRad gel dryer with Savant pump for 35 minutes at 80 °C and reaction products were visualized by exposure to a phosphorscreen and subsequent scanning with a phosphorImager.

2.8 Computer assisted sequence analysis

GeneTool 1.0 program (BioTools) was utilized for raw sequence handling and protein sequence alignment and oligonucleotide primer design. Similarities between *sco3550* predicted protein sequence and homologous proteins in the Sanger *S. coelicolor* database (http://www.sanger.ac.uk/Projects/S_coelicolor/) and other known proteins in the NCBI database (http://www.ncbi.nlm.nih.gov/BLAST) were uncovered using BLASTP searches. Potential RNA secondary structures at 30 °C,

representative of the standard laboratory growth temperature of *S. coelicolor*, were predicted with Mfold 2.3 located at the internet site http://www.ibc.wustl.edu/~zuker/rna/.

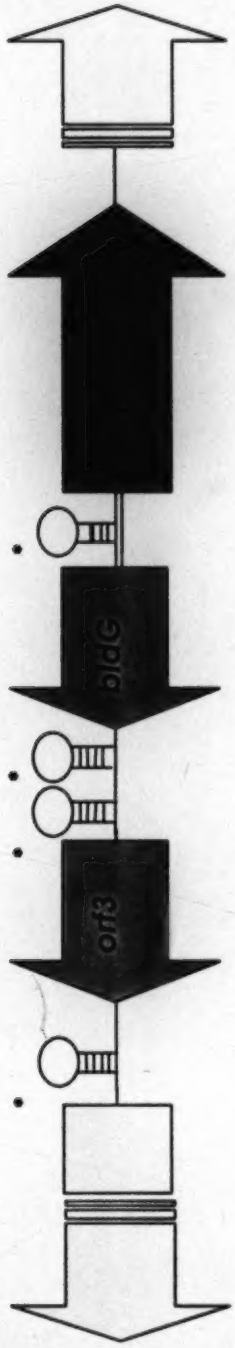
Chapter 3:

Results

3. Results

During the initial identification and sequencing of the *bldG* locus, a divergently transcribed open reading frame was identified that was predicted to encode a putative ATP-dependent RNA helicase and later designated *sco3550*, representing its position in the *S. coelicolor* genome (Bignell et al., 2000; Stoehr, 2001). RNA helicases are typically involved in regulatory processes by invoking RNA-dependent ATPase and ATP-dependent RNA unwinding activity upon short segments of RNA secondary structure. The role of RNA helicases appears to be as ubiquitous as the many routes of RNA metabolism in the cell and the destabilization, resolution and possible restructuring of RNA secondary structure is implicated in vastly separate processes. The close proximity of the genome project predicted GTG translational start site of SCO3550 to the *bldG* predicted translational start site and predicted promoters, as well as the divergent genomic positioning of the genes, raises the possibility of a coordinate control mechanism. Furthermore, this possiblity appears even greater with the identification of two promoters and possibly a third, controlling expression from the *bldG* locus in the short intergenic region between *bldG* and *sco3550* (Bignell, 2004, in press). Also, initial sequence analysis of the *bldG* locus identified three primary regions of RNA secondary structure via the Mfold prediction program that may require the controlling influence of an RNA helicase to maintain the complex mono- and polycistronic regulation seen at the locus (Bignell, Warawa *et al.* 2000) (Figure 3.1). An inverted repeat with potential stem-loop forming ability is present downstream of the *orf3* stop codon and may mediate

Figure 3.1 Schematic representation of the *bldG* locus. The gene of interest for this study, *sco3550,* is positioned divergently from *bldG* (*sco3549*) and *orf3* (*sco3548*). Potential regions of intramolecular RNA duplex formation or hairpin formation, predicted by Mfold analysis and possible targets for the activity of an RNA helicase, are indicated by (*).


aldG
ort3

transcriptional termination or exonuclease digestion of nascent transcripts. Also, two inverted repeats were identified in the *bldG- orf3* intergenic region that were predicted to be capable of forming hairpins and may play a role in the balance of polycistronic and *bldG* monocistronic transcripts at the locus (Bignell et al., 2000). Mfold analysis has also predicted the formation of RNA hairpins in the *bldG-sco3550* intergenic region that may affect the expression of *sco3550* and be involved in potential auto-regulation of the gene. It also cannot be discounted that there is potential for *bldG* and *sco3550* intermolecular RNA duplex formation due to transcript overlap that may affect gene expression and act as a candidate substrate for the action of a RNA helicase.

The *bldG* gene derives its name from its mutant phenotype; *bldG* mutants exhibit the absence of any aerial hyphal growth or sporulation. As well, *bldG* mutants show a near complete lack of pigmented antibiotic production, with the exception of faint pink cell-associated pigment most-likely representative of delayed undecylprodigiosin expression. Consequently, the *bldG* encoded anti-anti sigma factor is hypothesized to play a vital role in morphogenesis and physiological differentiation of the developing *S. coelicolor* colony. In this vein, it is also predicted that the *orf3* encoded anti-sigma factor is responsible for the regulation of a sigma factor or factors involved in differentiation and that BldG modulates this regulation and effectively plays a controlling role in the timing of development (Bignell et al., 2003). Finally, the genomic location of the putative RNA helicase, *sco3550*, suggests that the position of the gene may affect development through the regulation of the convoluted expression pattern observed from the *bldG* locus (Bignell et al., 2000). A coordinate control mechanism between the overlapping *bldG* and *sco3550* locus may take effect by a supercoiling-dependent

mechanism or steric restrictions imposed upon RNA polymerase holoenzymes involved in divergent expression. Alternatively, the *sco3550* encoded putative RNA helicase may function to regulate the expression of *bldG* and *orf3* through many means. Therefore, the ongoing objective of both this thesis project and the previous project of former MSc student, Julie Stoehr, was to define a link between *sco3550* and development in *S. coelicolor* and elucidate any involvement in the regulatory interplay of the *bldG* locus and the underlying *bld* gene architecture.

Prior to the onset of this thesis project, previous characterization of *sco3550* provided some tenuous links to the involvement of the putative RNA helicase in *S. coelicolor* development. In this respect, transcript analysis was attempted in order to examine the temporal profile of gene expression for *sco3550* within a developmental context, however, the transcript proved to be highly intractable to analysis. Owing to potential low abundance, instability and/or secondary structure of the transcript, northern analysis with DNA and RNA probes, S1 nuclease protection assays, primer extension and promoter probe analysis using both single and multi-copy expression were unable to identify the *sco3550* encoded transcript. Yet, semi-quantitative RT-PCR allowed detection of the mRNA transcript and revealed a biphasic mode of gene expression whereby peak levels of *sco3550* transcript would occur at the 15 hour vegetative growth time point and the 36 hour point of growth phase transition and onset of sporulation (Stoehr, 2001). This prompted speculation that the putative helicase may be playing a role in both vegetative and morphogenetic aspects of the complete life cycle. The use of RT-PCR also confirmed the low level expression of *sco3550* when in comparison with the internal control, the vegetative sigma factor *hrdB* transcript. This technique was also

used to localize the transcriptional start site of *sco3550* through primer walking and positioned it between the distal and proximal *bldG* promoters [at this stage only two *bldG* promoters had been identified (Bignell et al., 2000)], which is also in the region of the *S. coelicolor* genome project predicted GTG start codon for SCO3550. This together with the identification of a suitable Shine-Dalgarno ribosome binding site, strongly influenced the assignment of an in-frame ATG start codon, approximately 114 nt downstream of the genome project predicted GTG codon, as the *sco3550* translational start site. In the preliminary work of Julie Stoehr, mutagenesis of *sco3550* was also attempted in order to observe developmental consequences and examine the effect that a *sco3550* null mutation would have on the expression of the *bldG* locus. However, gene disruption was unsuccessful and it was hypothesized that this was due to its biphasic gene expression profile, with the putative helicase most likely playing an essential role in some portion of the cell cycle. The ability to disrupt *sco3550* in a *bldG* null mutant strain led to the final supposition that the putative helicase may act as a repressor of sporulation; thus, the *sco3550* single mutation is lethal because under these conditions development would no longer be repressed and the strain would be unable to cope with the resulting accelerated life cycle. On the other hand, a disruption of *sco3550* in a *bldG* sporulation deficient background would produce a strain that would not be able to differentiate even in the absence of the SCO3550 repressor (Stoehr, 2001).

The goal of this study was to further investigate the role of the *sco3550*-encoded putative RNA helicase in the developmental program of *S. coelicolor*. Gene expression profiles were reexamined and protein expression profiles were analyzed in the attempt to identify further links with development and the *bldG* locus. Also, mutagenesis was

performed utilizing the newly developed powerful technique of PCR-targeted mutagenesis in order to examine molecular and phenotypic consequences of *sco3550* gene disruption. The function of the putative RNA helicase *in vivo* was also more directly examined by confirming the predicted biochemical activity of recombinant SCO3550, in so far as the execution of *in vitro* RNA unwinding. Finally, the overexpression of a 6× histidine-tagged SCO3550 fusion protein in *S. coelicolor* and subsequent chemical crosslinking experiments were performed in an attempt to identify any possible protein-protein interactions between other cellular proteins and the putative RNA helicase, with the long-term goal to further clarify the cellular and metabolic context of the SCO3550 protein.

3.1 Sequence analysis of *sco3550*

BLAST searches (unfiltered) of the NCBI database revealed that SCO3550 has a high degree of homology with other predicted ATP-dependent RNA helicases and putative helicases from other actinomycetes. Furthermore, SCO3550 shared the highest identity (92%) and similarity (96%) with a putative RNA helicase in *Streptomyces avermitilis* (SAV4613). Interestingly, the entire sequence and synteny of the arrangement of *sco3550* and the *bldG* locus appears well-conserved in *S. avermitilis* and lends strength to the proposed link between *sco3550* and *bldG-orf3*. Other close homologs of SCO3550 were found in *Thermobifida fusca* and *Nocardia farcinica*, as well as other actinomycetes, however no function for the homologs has been characterized in these organisms. Subsequent alignment of the protein sequences shows conservation of RNA helicase motifs involved in ATP binding and hydrolysis, nucleic acid binding and

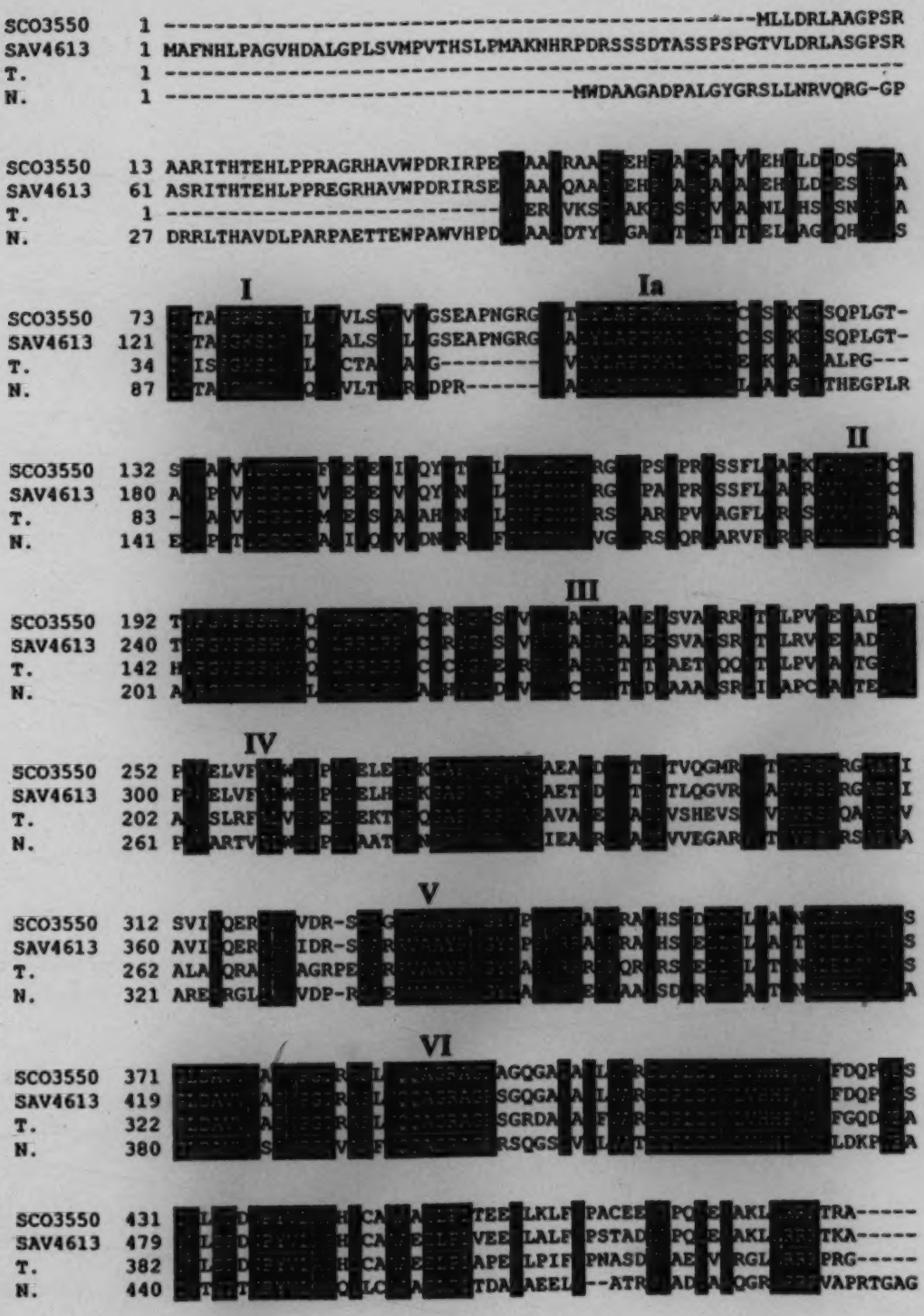
coordination of RNA unwinding (Figure 3.2). Like the helicase catalytic core, the aligned proteins appear to show some homology in the C-terminal domain. A unique domain was identified in SCO3550 and its homologs and was designated by the NCBI database as COG1205, a disinct helicase family with a unique C-terminal domain including a metal-binding cysteine cluster. Although no references could be found on the subject of COG1205, metal-binding cysteine clusters are often found in a paired conformation, such as [$CX_{1-4}CX_2CX_{1-8}C$] (personal communication, K. Colvin) and such a domain does reside at the extreme C-terminus of SCO3550. The potential for chelation of a metal ion adds a further degree of complexity to the regulation and action of SCO3550 and must be considered for a complete evaluation of the overlapping *sco3550/ bldG* locus.

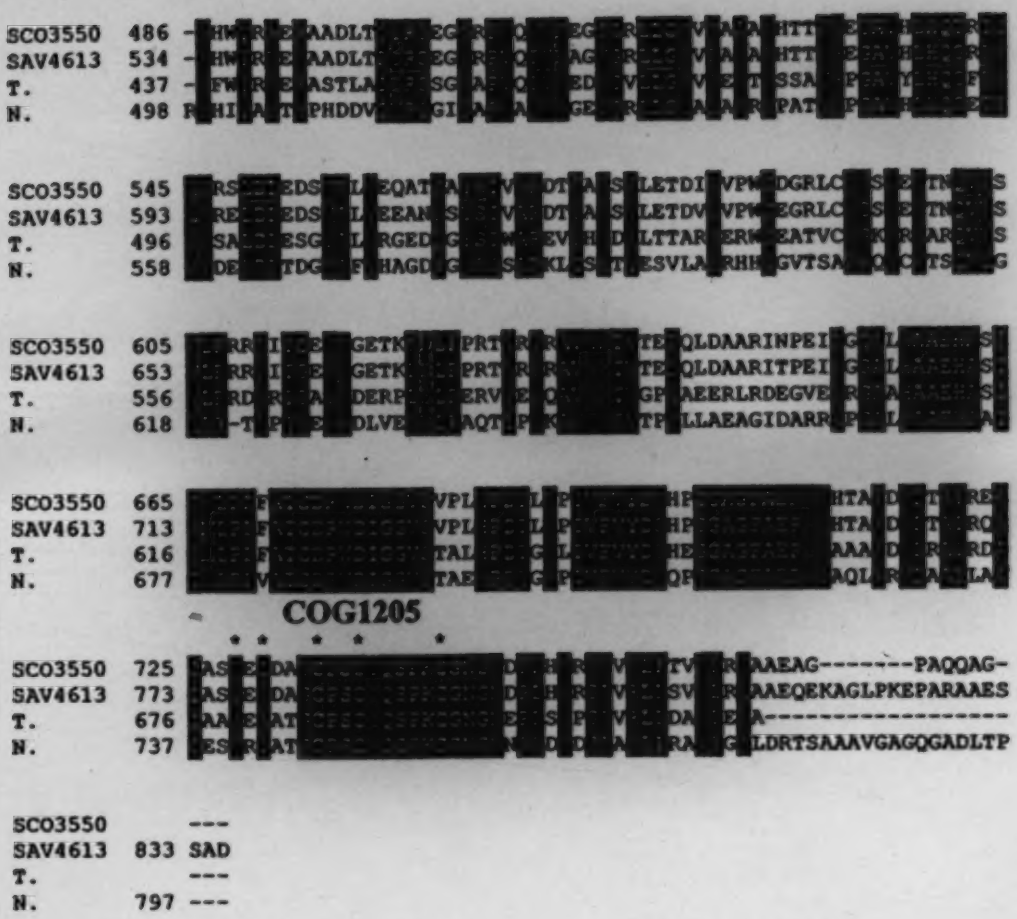
3.2 MBP-SCO3550 Overexpression and Purification in *E. coli*

Overexpression and purification of SCO3550 from *E. coli* was performed in order to isolate SCO3550 for use in the generation of poly-clonal antibodies, and preliminary *in vitro* enzyme assays to confirm putative helicase biochemical activity. For these studies, SCO3550 was expressed as a MBP fusion in BL21 (DE3), a protease deficient strain of *E. coli*, and the recombinant protein was purified by amylose resin chromatography.

Design of the overexpression construct involved the cloning of the coding sequence for *sco3550* into the pMAL-c2X vector downstream of the *malE* gene, encoding MBP, to create pAU330 (Table 2.4). The intervening sequence between the coding regions of *malE* and *sco3550* contains a manufacturer-engineered protease cleavage site, where the MBP tag can be removed from purified SCO3550 through

Figure 3.2 Protein sequence alignment of SCO3550 with putative RNA helicase homologs found in other organisms. The alignment was generated using the ClustalW Program and was displayed using the Boxshade Program (Version 3.2.1). SCO3550 was aligned with a sample of putative RNA helicases showing significant homology (>50% identity [ID]) from *Streptomyces avermitilis* (SAV4613; 92% ID, 96% similarity), *Thermobifida fusca* (T.; assession ZP00292213, 63% ID, 74% similarity) and *Nocardia farcinica* (N.; assession YP116571, 57% ID, 66%). Sequence boxshaded in black represents 100% consensus between the aligned sequences and sequence boxshaded in grey represents at least 50% consensus (<100%) between the sequences. Note that conserved RNA helicase motifs are indicated in the protein alignment: I (GKS), Ia (PTKALA), II (DExH), III (SAT), IV (LVFAL), V (LAGRVAAYRG), VI (QQAGRAGR). Also, the C-terminal domain-containing metal-binding cysteine cluster which represents a distinct family of helicases, is indicated as COG1205 {[$CX_{1-4}CX_nCX_{1-8}C$]} and cysteines that are predicted to be involved in chelation are indicated by (*).


SCO3550 1 ----------MLLDRLAAGPSR
SAV4613 1 MAFNHLPAGVHDALGPLSVMPVTHSLPMAKNHRPDRSSSDTASSPSPGTVLDRLASGPSR
T. 1 ----------
N. 1 ----------MWDAAGADPALGYGRSLLNRVQRG-GP
SCO3550 13 AARITHTEHLPPRAGRHAVWPDRIRPE
SAV4613 61 ASRITHTEHLPPREGRHAVWPDRIRSE
T. 1 ----------
N. 27 DRRLTHAVDLPARPAETTEWPAWVHPD
I
Ia
SCO3550 73
SAV4613 121
T. 34
N. 87
II
SCO3550 132
SAV4613 180
T. 83
N. 141
III
SCO3550 192
SAV4613 240
T. 142
N. 201
IV
SCO3550 252
SAV4613 300
T. 202
N. 261
V
SCO3550 312
SAV4613 360
T. 262
N. 321
VI
SCO3550 371
SAV4613 419
T. 322
N. 380
SCO3550 431
SAV4613 479
T. 382
N. 440

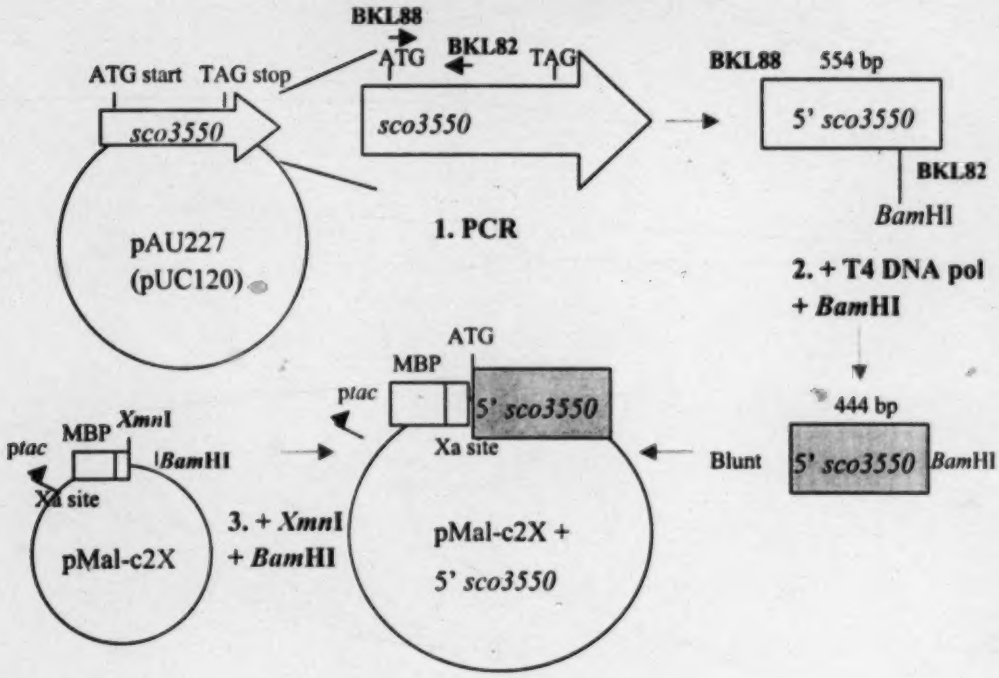

SCO3550 486
SAV4613 534
T. 437
N. 498
SCO3550 545
SAV4613 593
T. 496
N. 558
SCO3550 605
SAV4613 653
T. 556
N. 618
SCO3550 665
SAV4613 713
T. 616
N. 677
COG1205
SCO3550 725
SAV4613 773
T. 676
N. 737
SCO3550 ---
SAV4613 833 SAD
T. ---
N. 797 ---

Factor Xa protease cleavage. The initial cloning involved a two-part strategy because the 5' portion of *sco3550* required a blunt end at the predicted ATG start codon in order to be cloned into the pMAL-c2X vector in-frame with the MBP tag (Figure 3.3). This necessitated the use of a PCR primer that included sequence up to the ATG start codon but because of the large size of *sco3550*, PCR could not be utilized to amplify the entire gene. Consequently, a 560 bp PCR product was amplified using primers BKL88, which annealed at the predicted *sco3550* ATG start codon, and a downstream primer, BKL82, (Table 2.5) from the pAU227 template. The PCR product was first treated with T4 DNA polymerase to remove protruding 3' adenosines and was then digested with *BamHI* to remove approximately 100 bp from the BKL82 side of the fragment. The resulting blunt ended- *BamHI* digested fragment containing a 5' portion of *sco3550* was cloned into *XmnI/BamHI* digested pMAL-c2X vector and positive transformants were sequenced to ensure the integrity of the insert sequence. The remaining *sco3550* coding region (3' end) was prepared for insertion downstream of this small 5' fragment by an initial *NcoI* digestion of pAU227 followed by blunting of the resulting sequence overhang and the liberation of an approximate 2.2 kb fragment of *sco3550* from the linear and blunted pAU227 plasmid by *BamHI* digestion. pMAL-c2X containing the 5' *sco3550* fragment was prepared for ligation by *BglII* digestion followed by blunting, and then by *BamHI* digestion. Finally, the resulting *BamHI*/blunt 2.2 kb fragment containing the 3' *sco3550* sequence was cloned into the *BamHI*/blunt digested pMAL-c2X already containing the 5' end of *sco3550*.

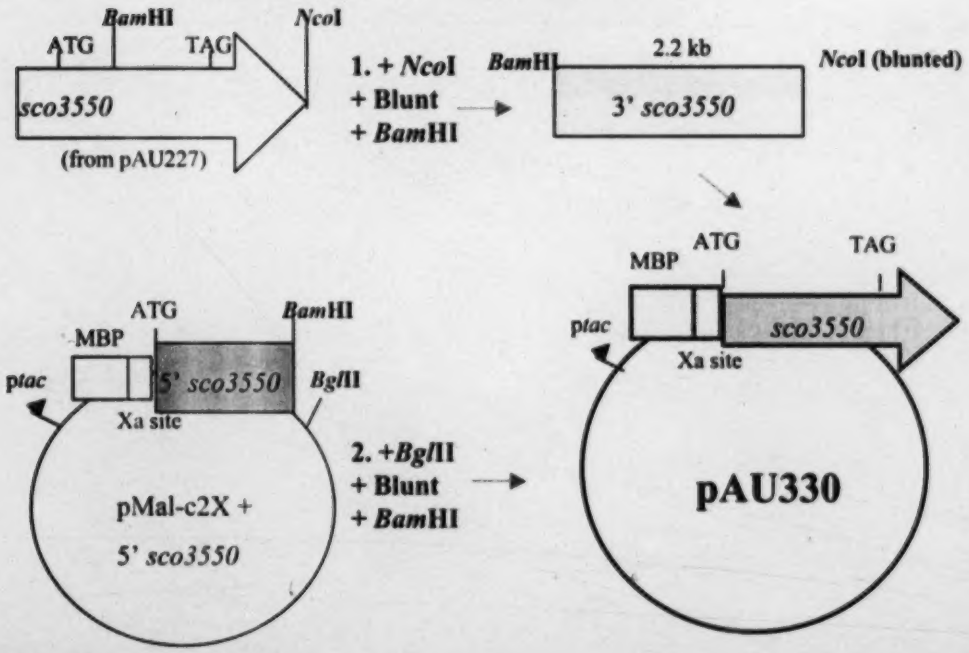
Expression of MBP-SCO3550 was performed in protease-deficient *E. coli* BL21 (DE3). In order to maximize the amount of soluble fusion protein present in cell extracts,

Figure 3.3 Schematic depiction of pAU330 construction. The coding region of *sco3550* and 352 bp downstream of the TAG stop codon was cloned into pMal-c2X, the *E. coli* MBP tag overexpression plasmid, to construct plasmid pAU330. The cloning strategy entailed a two part procedure as PCR amplification of the 5' region of the gene was necessary to engineer a blunt end at the start codon that was critical for cloning *sco3550* in frame with the MBP tag coding sequence beginning at the ATG start codon (Note a Factor Xa Protease target site is also indicated and located downstream of the MBP tag in pMal-c2X). Step 1 consisted of (1.) PCR amplification from the pAU227 template with primers BKL88 and BKL82, this was followed by (2.) T4 DNA polymerase blunting of the PCR fragment and a *Bam*HI digestion. The blunt/*Bam*HI 5' *sco3550* was ligated with the *Xmn*I/*Bam*HI digested pMal-c2X plasmid to form pMal-c2X+5' *sco3550*. Step 2. Consisted of cloning of 3' *sco3550* into pMal-c2X+5' *sco3550*. pAU227, containing the *sco3550* coding sequence as well as regions upstream and downstream of the gene, was (1.) digested with *Nco*I, which cleaves downstream of the TAG stop codon. The plasmid was then blunted and a 2.2 kb fragment containing 3' *sco3550* and sequence downstream of the TAG stop codon was liberated by a *Bam*HI digestion. (2.) pMal-c2X-5' *sco3550* was prepared by a *Bgl*II digestion, followed by blunting and a *Bam*HI digestion and the 3' *sco3550* fragment was ligated to the 5' *sco3550* in the pMal-c2X+5' *sco3550* plasmid through a *Bam*HI/blunt ligation. This two part cloning strategy yielded pAU330, an *E. coli* overexpression plasmid that imparts p*tac* mediated overexpression of the MBP-SCO3550 fusion protein.

Step 1: 5' *sco3550* cloning

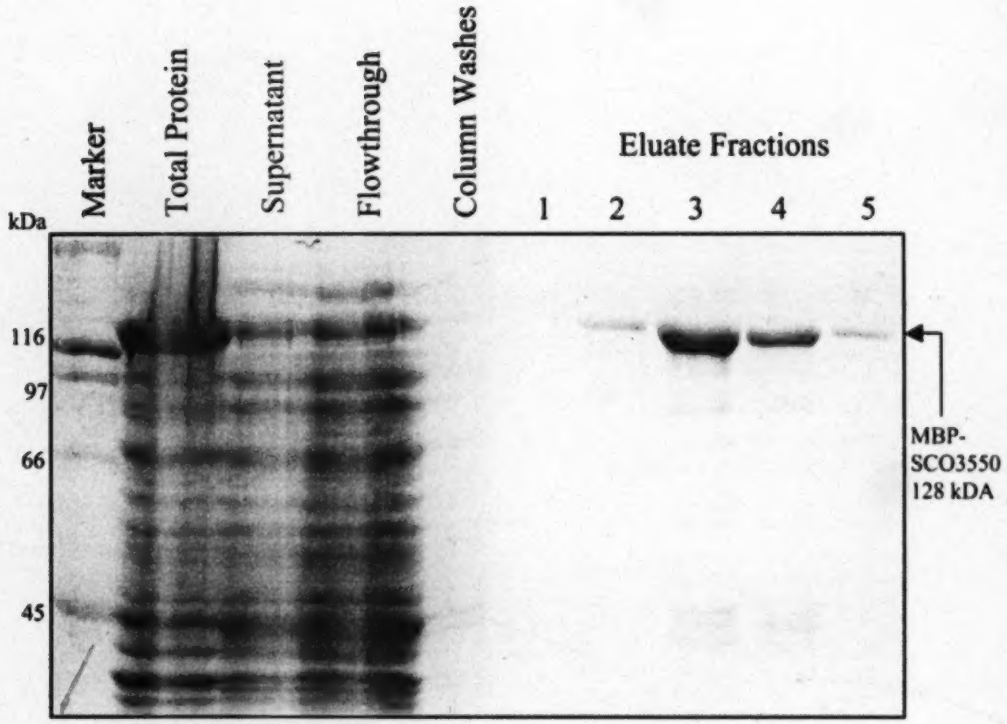


Step 2: 3' *sco3550* cloning



the overexpression strain was grown at 30 °C, rather than 37 °C, to the desired OD_{600nm} prior to induction of the p*tac* promoter by addition of IPTG to 0.3 mM. Again to prevent formation of insoluble aggregates of fusion protein, the induction was performed at room temperature for 6 hours rather than the standard 2 hour 37 °C induction suggested by the manufacturer. Cells were frozen overnight to promote rapid lysis and subsequently thawed, sonicated and diluted in column buffer prior to amylose resin chromotography. As expected, the noted modifications of the above procedure increased the percentage of soluble recombinant protein relative to insoluble from virtually negligible amounts to approaching 50% of the total crude extract concentration. Chromotography was performed through loading the diluted crude extract onto the column, collecting flowthrough for analysis, followed by several intervening washes to remove cellular contaminants from the column and reduce the likelihood of non-specifically bound proteins. Finally, recombinant protein was purified in several fractions by elution with maltose, a competitor for the amylose bnding of MBP. Fractions containing the greatest amount of recombinant protein also commonly exhibited contaminant proteins that were observed by 7.5 % SDS-PAGE as lower molecular weight bands on the gel (Figure 3.4). These contaminants were believed to either be degradative products of the fusion protein, which is an inherent difficulty in overexpression experiments, or endogenous maltose-binding proteins from the host that are co-purified with the fusion protein. However, these faint contaminant bands were deemed to be of little concern to the present studies as gel-purified fusion protein was used for the generation of poly-clonal antibodies and only fractions bereft of any visible contamination were used for the *in vitro* helicase assays (see later, Section 3.7).

Figure 3.4 Overexpression and purification of MBP-SCO3550. Fusion protein MBP-SCO3550 was overexpressed in *E. coli* BL21 (DE3) from the p*tac* promoter by addition of 0.3 mM IPTG inducer. Prior to lysis, a 20 μL aliquot of the cell pellet was boiled and combined with 5 μL 2xSDS-PAGE loading buffer (Total Protein). This is compared to the amount of soluble fusion protein in a 20 μL aliquot of the supernatant, which was prepared after sonication of the remaining cell pellet and subsequent pelleting of cellular debris (Supernatant). Total column flowthough (0.5%; Flowthrough), column washes (0.5%; Column Washes) and elutions with column buffer containing 10 mM maltose (10%; Eluted Fractions) were resolved by 7.5 % SDS-PAGE. The gel was stained with Coomassie Brilliant Blue. Band sizes are indicated for the molecular weight marker SDS-PAGE Broad Range Standards (BioRad), and the position and expected size of the MBP-SCO3550 fusion protein is also indicated.


Marker
Total Protein
Supernatant
Flowthrough
Column Washes
Eluate Fractions
1
2
3
4
5
kDa
116
97
66
45
MBP-
SCO3550
128 kDA

3.3 Protein Analysis

3.3.1 Western analysis of SCO3550

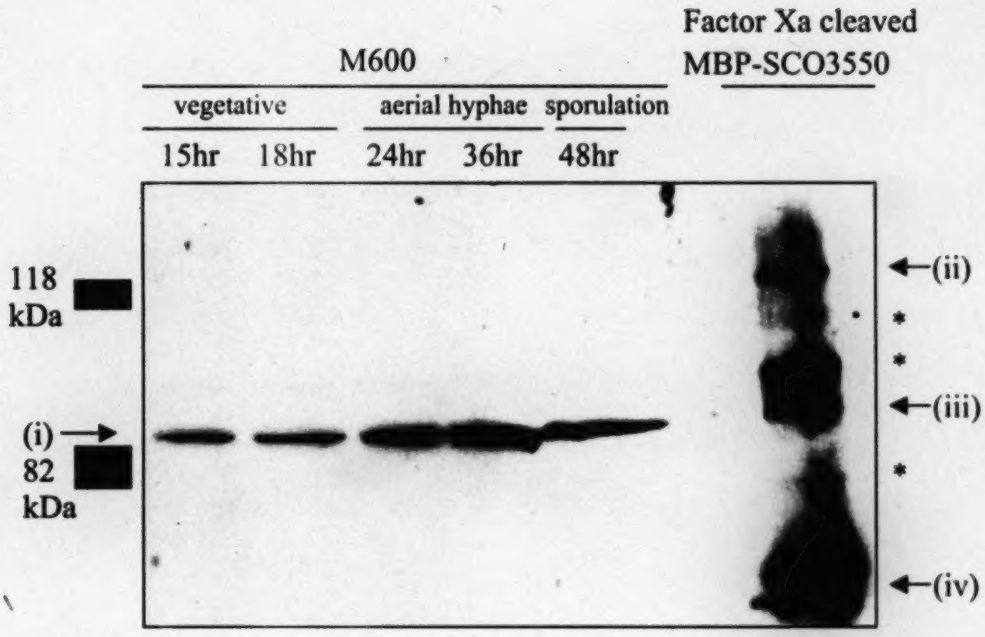
In an effort to observe SCO3550 expression at all stages of colony development, wild-type *S. coelicolor* M600 was grown on the surface of R2YE agar plates and samples were isolated at various times. Total cell extracts were isolated and western analysis was performed on at least three separate sets of time course samples using the anti-SCO3550 serum raised against MBP-SCO3550 purified from *E. coli*. It was also hoped that these experiments would lend validity to the biphasic temporal profile of gene expression observed by Stoehr, where *sco3550* transcript was detected throughout the life cycle with peak levels of gene expression occurring at 15 and 36 hours of growth (Stoehr, 2001). Forty-fifty microgram aliquots of total cell protein were separated by 7.5 % SDS-PAGE and transferred to a PVDF membrane. It should be noted that 15 and 18 hour time points correspond to vegetative growth while the 24 and 36 hour points include the initial growth stage transition manifested as the appearance of pigmented antibiotic production and erection of aerial hyphae, which culminates with complete hyphal sporulation at the final 48 hour time point. Initially, Factor Xa cleaved fusion protein acted as a positive control for SCO3550 protein, however, many putative degradative products or cross-reacting contaminants from the protein isolation experiment were also detected by western analysis and confounded the comparison between the cleaved fusion protein and the correct cell extract hybridizing band. These cross-reacting products also hindered the attempt to quantify the protein expression levels of SCO3550 relative to a standard curve of varying amounts of FactorXa cleaved fusion protein. For the primary antibody hybridization, the membranes were incubated with anti-MBP-SCO3550 serum (rabbit

IF6) at a dilution of 1 in 2000 and a secondary antibody dilution of 1 in 10 000. Secondary antibodies were covalently linked to horseradish-peroxidase and the membrane was processed through a brief incubation with chemiluminescent reagents and immuno-hybridizing bands were detected by short exposures to film. Preliminary results indicated that SCO3550 was expressed as a single isoform approximately 84-88 kDa in size, which was reflective of the predicted 86 kDa size calculated by analyzing the deduced amino acid content of the protein (Figure 3.5). Expression appears to transpire throughout growth with lowest levels occurring at the 15 hour time point and then gradually increasing to peak levels at 24 hours and then returning to lower levels at 48 hours. The exhibited temporal profile of protein expression does not negate the possibility of both a vegetative and developmental function for the protein; yet, the protein profile does not conform to the RT-PCR detected biphasic gene expression pattern. Of course, gene expression does not always serve as a determinant of protein expression especially at a complex gene locus where a coordinate control mechanism may influence transcription of *sco3550* as a means to regulate the expression of *bldG* and *orf3*. It was unclear at this point if the putative helicase protein was active and whether it might perform the same role throughout the time course of protein expression and play a role in *S. coelicolor* development.

3.3.2 Western analysis of SCO3550 using affinity-purified antibodies

Several problems still existed in the preliminary western hybridization experiments including the failure of the Factor Xa cleaved fusion protein positive control to definitively identify the hybridizing band in *S. coelicolor* cell extracts as SCO3550.

Figure 3.5 Western analysis of SCO3550 protein by 7.5 % SDS-PAGE. Crude extracts from wild-type *S. coelicolor* strain M600 were isolated at varying time points corresponding to vegetative growth and morphological differentiation (as indicated) and 50 μg of total cell protein from each sample was separated by 7.5 % SDS-PAGE. Upon transfer to a PVDF membrane, the samples were probed with anti-SCO3550 polyclonal antibodies at a 1:2000 dilution followed by a secondary hybridization with horseradish peroxidase-conjugated secondary antibody at a dilution of 1:10 000. Detection was performed by a 1 minute incubation with Enhanced Perkin Elmer chemiluminescence reagents and exposure to Kodak film for 10-15 seconds. Kaleidoscope pre-stained standards were used as a general size marker to localize the size of the detected bands and to monitor the efficacy of electroblot transfer of proteins to the PVDF membrane (not shown). Factor Xa cleaved SCO3550 recombinant protein was used as a positive control. The predicted SCO3550 band from the cell extracts is indicated (i), as well as uncleaved fusion protein (ii), cleaved SCO3550 (iii) and cleaved MBP (iv), while possible non-specific degradative products of the protease cleavage or contaminating cross-reacting bands are also indicated (*).


Factor Xa cleaved
MBP-SCO3550
M600
vegetative
aerial hyphae
sporulation
15hr
18hr
24hr
36hr
48hr
118
kDa
(i)
82
kDa
(ii)
(iii)
(iv)

Identification of the correct band corresponding to SCO3550 was also complicated by the lack of a true negative control as a *sco3550* null mutant had yet to be sucessfully constructed. Also, a longer exposure of the film revealed that the prepared anti-serum had very low specificity and a large sum of potentially cross-reacting and non-specific hybridizations could be visualized, some of which obscured the putative SCO3550 band. It was a possibility at this stage of analysis that one of these minor hybridizing bands was the true SCO3550 and that other SCO3550 isoforms might exist. In an effort to purge the cleared serum of antibodies that lacked specificity to the SCO3550 recombinant protein and may have been contributing to the high background present on the previously described western experiments, a procedure to affinity-purifiy MBP-SCO3550 specific antibodies was undertaken. Approximately 10 μg of affinity chromatography purified MBP-SCO3550 was separated in duplicate by 7.5 % SDS-PAGE and transferred to a PVDF membrane. Subsequently, one half of the membrane was stained with Ponceau S dye to correctly position the fusion protein relative to the kaleidoscope pre-stained molecular weight standard, and the approximated site was then excised from the unstained portion of the membrane as a 0.5 x 1 cm PVDF membrane piece. The membrane piece was then blocked with 10% skim milk, and incubated overnight with complete anti-MBP-SCO3550 anti-serum diluted in 1xPBS buffer. Hybridization was followed by a series of washes and poly-clonal antibodies specific to MBP-SCO3550 were then eluted by a low pH glycine treatment and neutralized to avoid acid hydrolysis. Affinity-purified antibody titer and specificity were examined by western analysis and a dramatic reduction in background hybridization was observed. Interestingly, western analysis of SCO3550, when resolved on a 7.5 % SDS-PAG and probed with a small

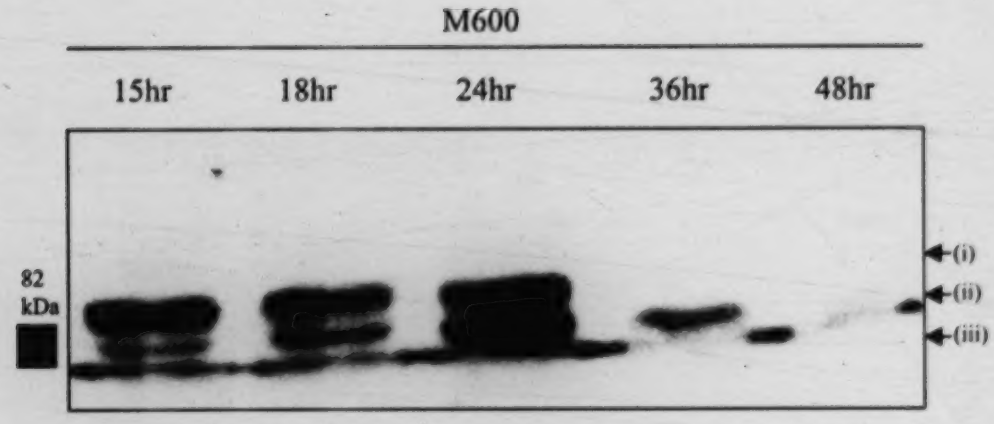
amount of affinity-purified antibody with a short exposure time, revealed a potential doublet where previous western experiments had showed a single band representative of SCO3550 protein. Western analysis was therefore performed on samples from at least three separate time courses that had been separated by 6 % SDS-PAGE and the potential doublet was resolved as two distinct bands by immunodetection: two representative figures are shown (Figure 3.6). This doublet may represent two isoforms of native SCO3550 or one of the bands may be a cross-reacting species, which was difficult to determine at this time due to the unavailability of a *sco3550* null mutant as a negative control (See later, Section 3.4 and 3.5). It is interesting to note that a protocol of affinity-purification of antibodies in combination with various percentages of polyacrylamide used in SDS-PAGE resolved the preconcieved SCO3550 single band as a doublet. On average, the proteome is predicted to represent 1.2 proteins per gene in *S. coelicolor* (Hesketh et al., 2002a), therefore variations upon this strategy of western analysis may be useful for proteome characterization.

3.4 Mutational studies of *sco3550*

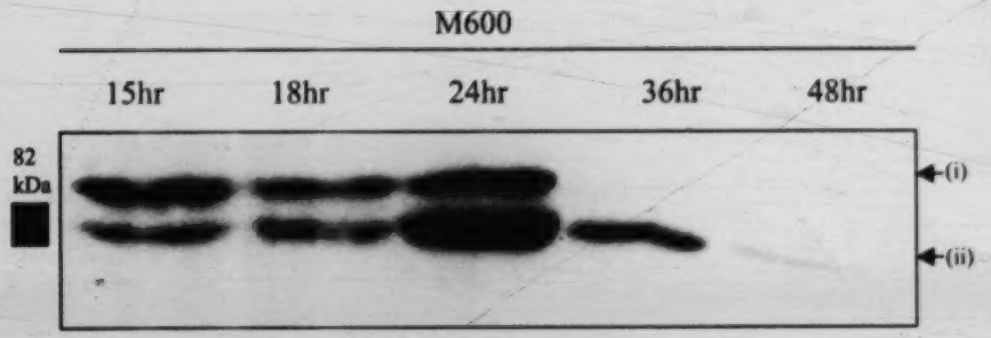
Mutagenesis studies were performed in an effort to construct a negative control for use in the identification of SCO3550 hybridizing bands by western analysis and to examine the effect of a *sco3550* gene deletion upon physiological and morphological differentiation in *S. coelicolor*. Previous experimentation by J. Stoehr had shown that *sco3550* is extremely recalcitrant to mutagenesis. In the experiments by Stoehr, a 690 bp *sco3550* internal DNA fragment was cloned upstream of a thiostrepton resistance gene to create pAU226, and a single cross-over disruption strategy was performed where

Figure 3.6 Western analysis of SCO3550 using affinity-purified antibodies (6 % SDS-PAGE). Crude extracts were isolated at various stages of colony development from wild-type *S. coelicolor* M600 grown on the surface of R2YE and separated by 6% SDS-PAGE. Samples were transferred to a PVDF membrane and probed with at least 10 µL of affinity-purified anti-SCO3550 antibodies, which had been previously titered through a series of standard western hybridization experiments. The primary hybridization was followed by the addition of Horseradish peroxidase-conjugated secondary antibody at a dilution of 1: 10 000. Detection of hybridized bands was performed by exposure of the membrane to a 1 minute incubation with Enhanced Plus (Perkin Elmer) chemiluminescence reagents and film exposure for 2 minutes. Note that Kaleidoscope pre-stained standards were used as a size marker for SDS-PAGE and western hybridization. (A) and (B) are representative western blots. (i) Initially assumed to be a cross-reacting band but was later speculated as a larger isoform of SCO3550 (see Section 3.5: Figure 3.15B and 3.16 are either partial or complete reproductions of Figure 3.6 B). (ii) The SCO3550 specific reacting band, later believed to be the smaller isoform of the protein. (iii) Indicates the dye front of the 6 % SDS-PAGE.

(A)


M600
15hr
18hr
24hr
36hr
48hr
82
kDa
(i)
(ii)
(iii)

(B)


M600
15hr
18hr
24hr
36hr
48hr
82
kDa
(i)
(ii)

homology between the native gene and the internal DNA fragment would promote recombination. In this manner, the *sco3550* ORF could be disrupted by the internal 690 bp fragment and a thiostrepton resistance gene and potential mutants could be selected by thiostrepton antibiotic resistance. Throughout repeated rounds of experimentation, thiostrepton resistant strains of *S. coelicolor* were not observed and therefore, it was presumed that *sco3550* was potentially immune to disruption and played some essential role in the life cycle of *S. coelicolor*. With only circumstantial evidence, it was difficult to conclude that *sco3550* was truly an essential gene or if mutagenesis suffered due to technical reasons, such as, a short region of homology in the disruption cassette, therefore further mutagenesis studies were undertaken.

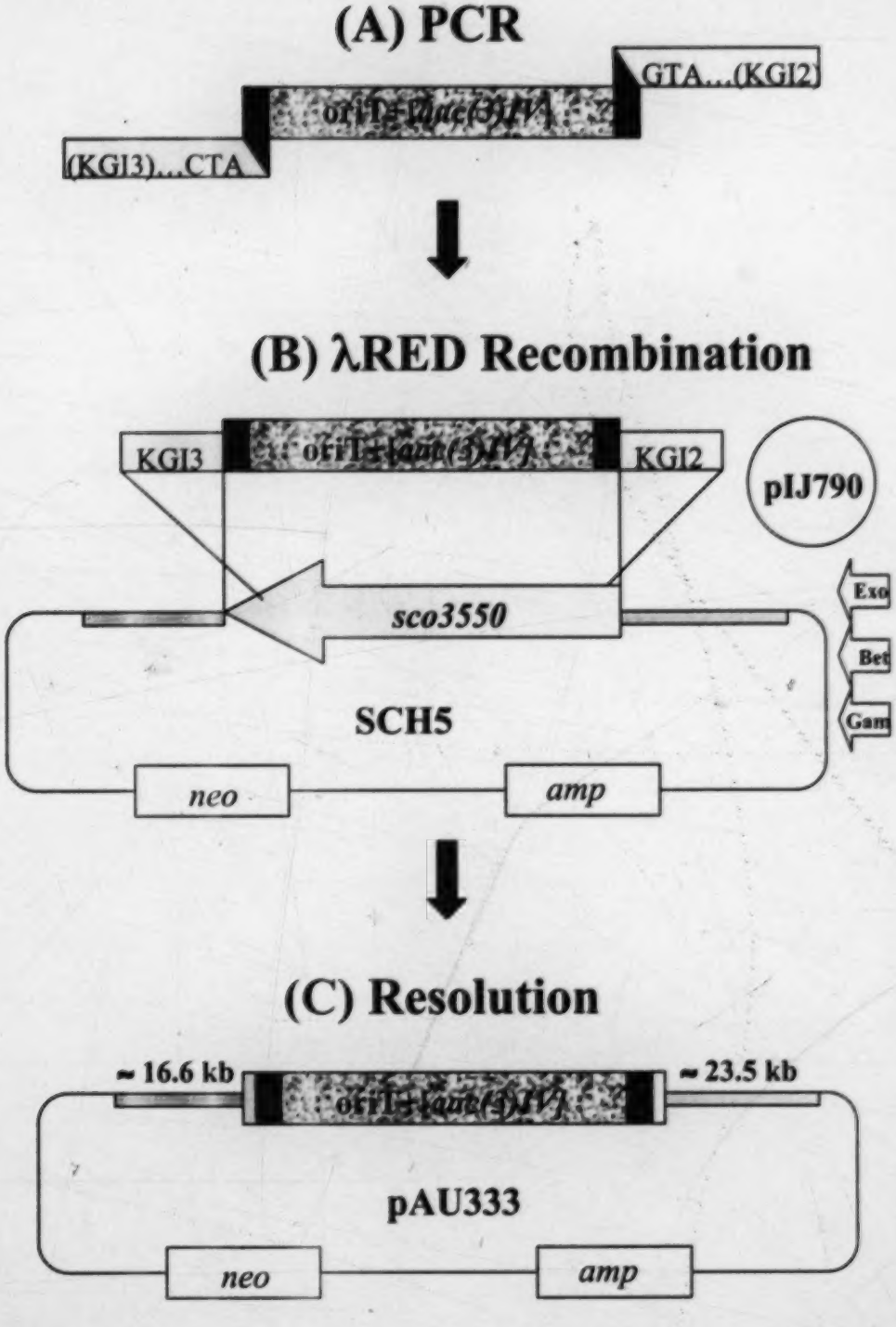
3.4.1 Deletion of *sco3550* in *S. coelicolor*

Construction of a *sco3550* mutant was performed utilizing a PCR targeting-based gene disruption protocol that was adapted for use in *S. coelicolor* and was designated Redirect© technology by the developers (Gust et al., 2003). The technique involves disruption in *E. coli* of a *S. coelicolor* cosmid with a linear DNA fragment consisting of an antibiotic resistance cassette, an *oriT* (conjugal origin of transfer) and flanked by short stretches of sequence homologous to intergenic sequence immediately upstream and downstream of the target gene. The disrupted cosmid is then isolated and transferred via conjugation from *E. coli* to a wild-type strain of *S. coelicolor*. A double-cross over that results in a replacement of the wild-type gene with the targeted antibiotic resistance cassette can be identified by the presence of the disrupting antibiotic resistance cassette and the loss of the cosmid specific antibiotic resistance marker. The power of the

technique lies in the final recombination stage, where rather than relying on a small region of homology to promote recombination, the disrupted gene is flanked by up to 20 kb of cosmid sequence homologous to the *S. coelicolor* chromosome and thereby facilitates a higher frequency of successful recombination events.

Figure 3.7 shows a schematic diagram detailing the construction of the apramycin resistance cassette [*aac(3)IV*]-disrupted SCH5 cosmid that was subsequently used for deletion of *sco3550* in the *S. coelicolor* chromosome. PCR amplification of the extended resistance cassette entailed the design of primers KGI2 and KGI3, which contained 39 nucleotides of homology to sequence upstream and downstream, respectively, of the *sco3550* start and stop codons (see Section 2.3.2). The PCR template was a gel-purified *Eco*RI/*Hind*III, 1384 bp apramycin resistance cassette [*aac(3)IV*, *oriT*] from pIJ773. PCR targeting of the *S. coelicolor* cosmid involved electroporation of the purified PCR product into electrocompetent *E. coli* BW25113 containing both pIJ790 (λRED recombination plasmid) and the SCH5 cosmid. L-arabinose (added to 10 mM) was used in the preparation of the electrocompetent BW25113 containing pIJ790 and SCH5 in order to induce expression of λRED proteins Redα (*exo*), Redβ (*bet*) and Redγ (*gam*) (Table 2.4). Transformants were plated on LB medium containing ampicillin and kanamycin to select for the SCH5 cosmid, and apramycin to select for replacement of wild-type *sco3550* from SCH5 with the apramycin disruption cassette. Furthermore, transformants were grown at 37 °C to promote the loss of pIJ790 and large colonies were chosen in order to select transformants that most likely contained a large number of mutagenised cosmids or multiple integrants. Confirmation of effective PCR-targeting of

Figure 3.7 Schematic protocol for PCR-targetting of *sco3550* deletion in cosmid SCH5. (A) The *sco3550* [*aac(3)IV+oriT*] gene replacement cassette was generated by PCR amplification with primers KGI2 and KGI3. Primers contained 19 (KGI3) and 20 nt (KGI2) regions homologous to the disruption cassette. The short regions of homology are indicated by the black triangular overlap between the schematic primers and apramycin resistance cassette. Primers also contained 39 nt extension homologous to *sco3550* intergenic sequence as indicated by the light boxes. (B) The *sco3550* gene replacement cassette was then used to transform electrocompetent *E. coli* BW25113/pIJ790 in the presence of arabinose in order to stimulate the expression of the λRED recombination protein (indicated by the white arrows). λRED-assisted recombination of the *sco3550* gene replacement cassette occurred by a double cross-over between regions of sequence homology residing in the flanking regions of the cassette and the SCH5 cosmid. (C) Upon resolution of the recombination event, transformants containing pAU333, or correct replacement of *sco3550* in the SCH5 cosmid with [*aac(3)IV+oriT*], were selected by apramycin resistance and confirmed by pAU333 isolation and restriction digestion analysis. Note that regions of wild-type *S. coelicolor* DNA in the cosmid are indicated by the thick grey bars and cosmid SCH5 specific sequence is indicated as the thin black line. Also, the white boxes in SCH5 and its *sco3550*-disrupted derivative, pAU333, represent the kanamycin-resistance gene (*neo*) and ampicillin-resistant gene (*amp*).


(A) PCR
GTA...(KGI2)
ori1+aac(3)IV
(KGI3)...CTA
(B) λRED Recombination
KGI3
ori1+aac(3)IV
KGI2
pIJ790
sco3550
Exo
Bet
Gam
SCH5
neo
amp
(C) Resolution
~ 16.6 kb
ori1+aac(3)IV
~ 23.5 kb
pAU333
neo
amp

the SCH5 cosmid was performed by cosmid DNA isolation and restriction digestion analysis with *Sac*I (Figure 3.8). All transformants were shown to contain correctly mutagenised cosmids (designated pAU333) and a sample from one was selected for transformation into electrocompetent *E.coli* ET12567/pUZ8002, followed by conjugation into *S.coelicolor* M600 spores. Potential exconjugants were grown on MS medium containing nalidixic acid, to select against further growth of *E. coli*, and apramycin to select for exconjugants containing the apramycin disruption cassette either as the result of a single or double cross-over into the chromosome (Figure 3.9). Note that the SCH5 cosmid does not contain a suitable origin of replication for propagation in *S. coelicolor* and exconjugants obtained from double cross-over events are expected to be free of cosmid DNA. Subsequently, exconjugants were replica-plated onto Difco nutrient agar (DNA), which allows fast, non-sporulating growth, containing nalidixic acid and apramycin, with and without the addition of kanamycin. Single cross-over exconjugants were presumed to be apramycin^R and kanamycin^R due to complete cosmid integration into the genome, whereas, double cross-over exconjugants were apramycin^R and kanamycin^S and indicative of a potential *sco3550* disruption. In these strains, the apramycin resistance cassette would replace the wild-type copy of *sco3550* in the chromosome and the cosmid, that now contains the wild-type copy of the gene, would be cleared from the strain because of an inability to replicate. Thirty-six kanamycin^S clones were streaked for single-colony isolates on MS agar containing nalidixic acid and apramycin, and kanamycin sensitivity was confirmed by replica-plating onto DNA medium containing nalidixic acid and apramycin, with and without the addition of kanamycin.

Figure 3.8 Confirmation of *sco3550* gene replacement in pAU333. Cosmid SCH5 and potential pAU333 isolates were digested with *Sac*I and examined by 1.0 % TBE agarose gel electrophoresis. The loss of a 3496 bp band of a doublet and gain of a 750 and 850 bp bands, relative to wild-type SCH5, in all potential pAU333 isolates indicated that *sco3550* gene replacement with [*aac(3)IV+oriT*] had occurred in all 6 transformants tested.

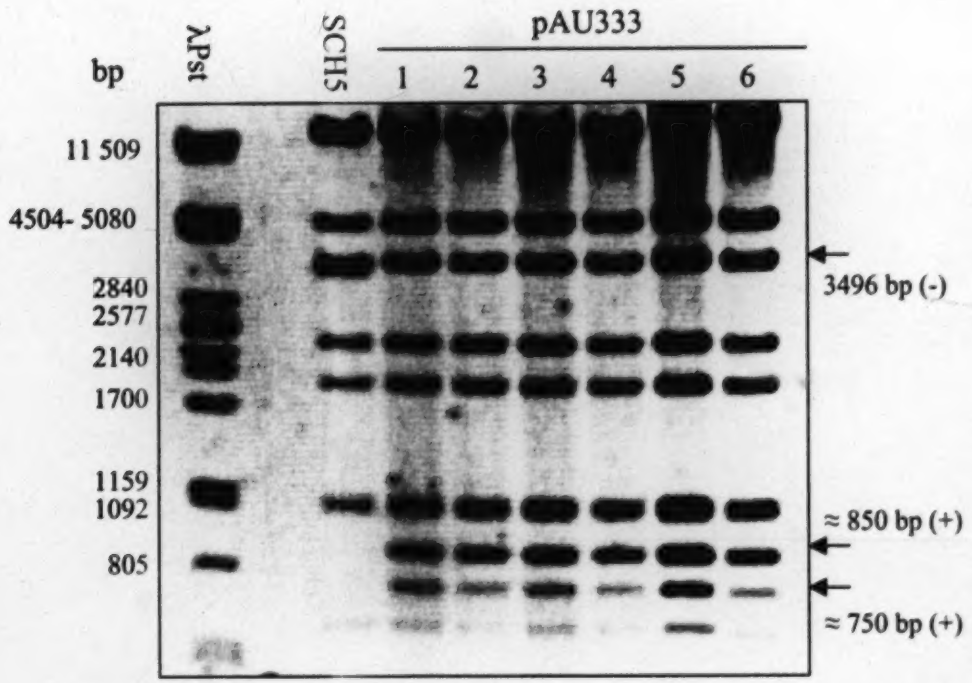
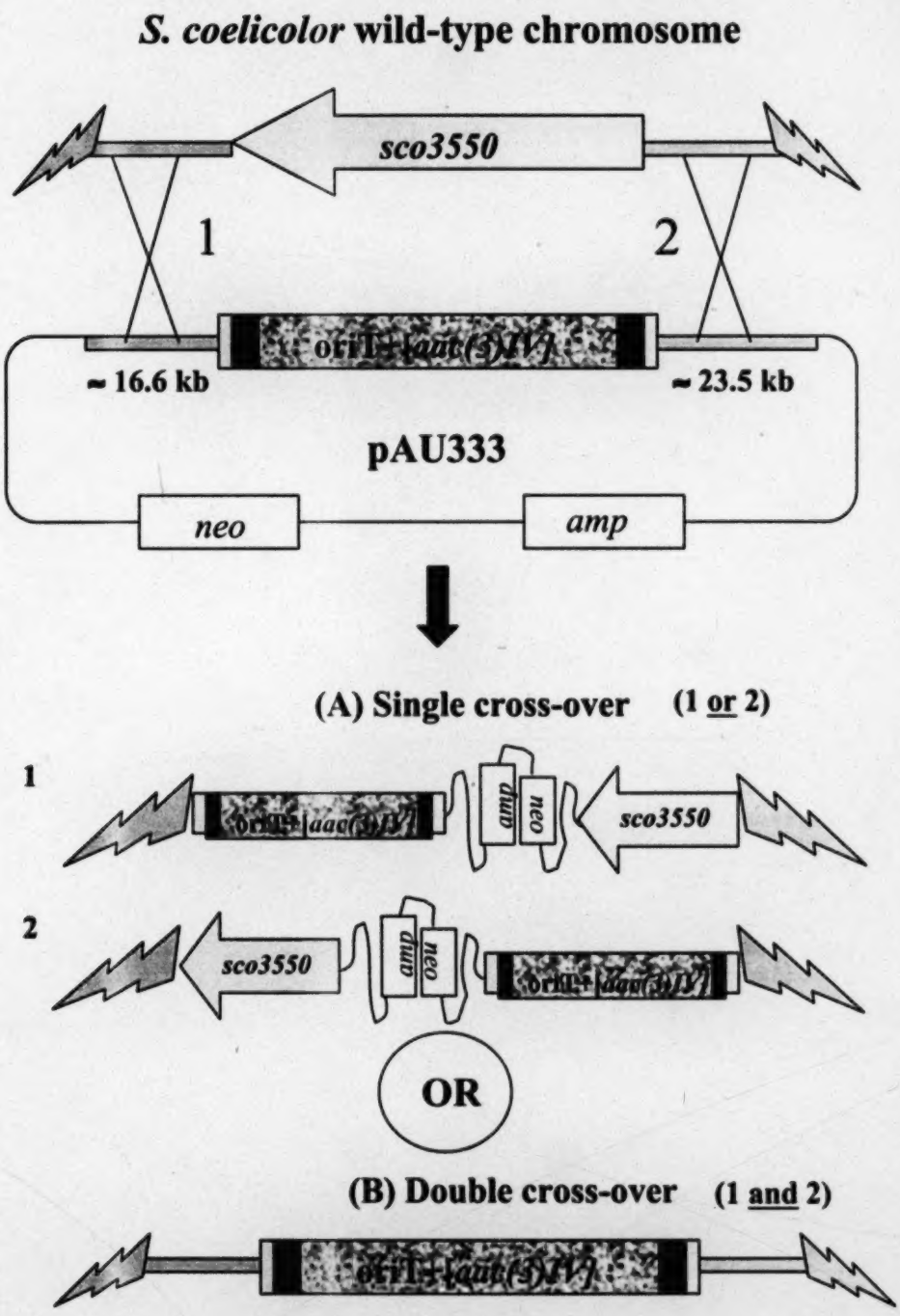

bp
λPst
SCH5
pAU333
1
2
3
4
5
6
11 509
4504- 5080
2840
2577
2140
1700
1159
1092
805
3496 bp (-)
≈ 850 bp (+)
≈ 750 bp (+)

Figure 3.9 Schematic protocol for replacement of *sco3550* in the *S. coelicolor* chromosome. The *sco3550* disrupted cosmid, pAU333, was transferred by conjugation from *E. coli* ET12567/ pUZ8002 into wild-type *S. coelicolor* and apramycin-resistant ex-conjugants were screened for kanamycin sensitivity. (A) is represenative of a possible fate of a single cross-over recombination event where both the disruption cassette and wild-type gene would reside in the chromosome and the strain would exhibit both apramycin and kanamycin resistance. (B) Double cross-over events would yield apramycin resistance and kanamycin sensitivity due to the replacement of wild-type *sco3550* with the disruption cassette. Note that regions of the chromosome and pAU333 sequence that potentially could undergo a recombination event are indicated by the thick grey bars, while chromosomal sequences outside the cosmid regions of homology are indicated by the grey lightning bolts. Cosmid SCH5 specific sequence is indicated by the thin black line and white boxes represent the kanamycin-resistance gene (*neo*) and ampicillin-resistant gene (*amp*) in the cosmid sequence.


S. coelicolor wild-type chromosome
sco3550
1
2
oriT+aac(3)IV
~ 16.6 kb
~ 23.5 kb
pAU333
neo
amp
(A) Single cross-over (1 or 2)
1
oriT+aac(3)IV
amp
neo
sco3550
2
sco3550
amp
neo
oriT+aac(3)IV
OR
(B) Double cross-over (1 and 2)
oriT+aac(3)IV

Consequently, all apramycin^R kanamycin^S clones continued to exhibit kanamycin sensitivity and 8 clones were selected for further analysis.

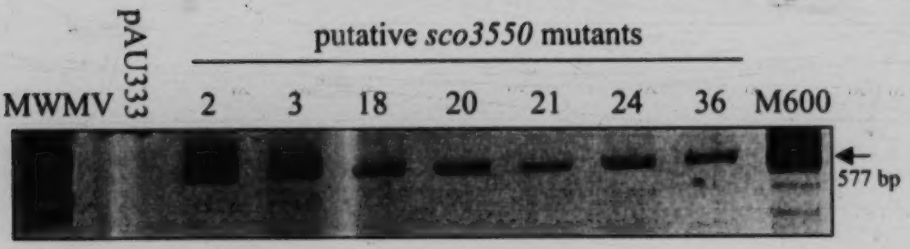
Prior to the initiation of experiments to confirm *sco3550* disruption, putative disruption strains were analyzed for phenotypic variance on R2YE medium and minimal medium containing glucose. However, putative *sco3550* disruption mutants showed no discernible phenotypic differences from wild-type strains of *S. coelicolor* (data not shown). One possibility for the lack of a visible phenotype was that due to the multi-genomic nature of *Streptomyces* vegetative mycelia, apramycin^R kanamycin^S clones may still contain a small proportion of wild-type chromosomes or chromosomes containing single cross-over disrupted cosmid integrants. These minor contributions of wild-type chromosomes could account for the wild-type phenotype observed in the putative mutants without providing the necessary levels of kanamycin resistance to manifest at the phenotypic level. To ensure that putative mutant clones arose from a single disrupted genome progenitor, spore stocks were created prior to further phenotypic and molecular characterization. Instances of sporulation in *Streptomyces* species constitute the only genuine cell division in developing mycelium, therefore, a spore will contain a single chromosome and growth arising from a single spore should not contain a mixture of distinct chromosomal species (Flardh et al., 1999). Spore stocks were created for 7 putative SCO3550 deletion mutants and spore stocks were replica-plated onto R2YE medium containing apramycin, with and without the addition of kanamycin. Surprisingly, all clones displayed apramycin^R and kanamycin^S and were assumed to be, under these conditions, putative disruption mutants possessing a pure complement of double cross-over disrupted chromosomes. A single clone from each precursor strain

was then grown for phenotypic analysis and tested by Southern analysis and chromosomal PCR for the appropriate disruption. Once again, clones revealed no phenotypic difference from that of wild-type *S. coelicolor* but it is always possible that null mutations may not manifest at a visible phenotypic level (such as a defect in morphological differentiation or pigmented antibiotic production), especially in *Streptomyces*, where there is the potential for molecular redundancy. Also, some phenotypes must be flushed out in the case of sensitivities to various environmental conditions, such as osmotic stress.

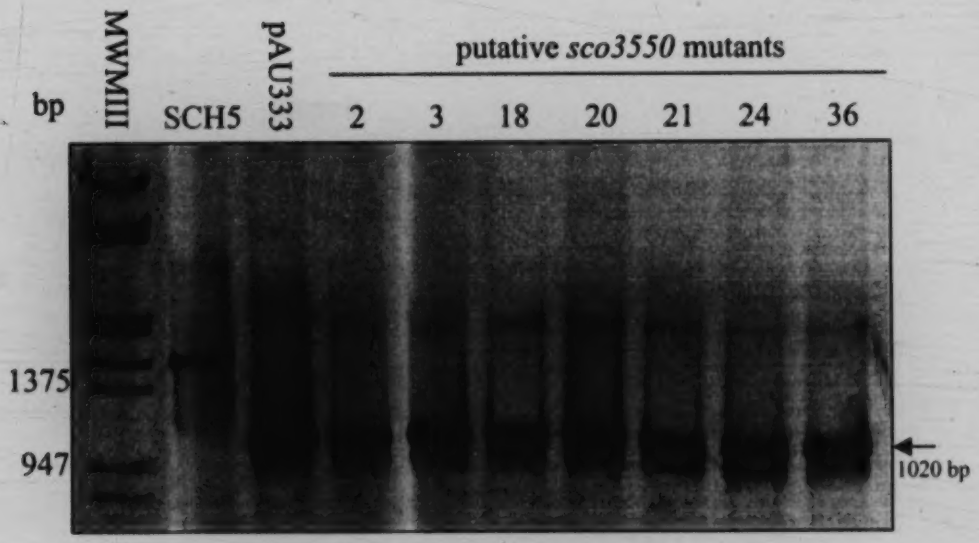
Before an extensive survey of growth characteristics of the putative mutants was attempted, molecular confirmation of the *sco3550* disruption mutation was performed using chromosomal PCR. Chromosomal DNA was isolated for the putative mutant clones, DNA was quantified, denatured and PCR performed with primers KGI4 and KGI5 to amplify an internal fragment of *sco3550*. Ideally, primers would be used that flank start and stop codons of the gene in question and thus, would exhibit a size difference between the PCR amplified wild-type gene and the apramycin cassette-disrupted gene, however the *sco3550* ORF is too large for reliable PCR amplification (> 2 kb). Therefore, chromosomal PCR was attempted and separation of the amplified DNA by 5 % PAGE revealed an amplified product of the expected size, suggesting that all of the putative mutant strains contained an intact copy of *sco3550* (Figure 3.10 A). Chromosomal PCR was also performed utilizing primers JST5, which annealed upstream of the *sco3550* ATG start codon, and KGI6, which anneals within the apramycin [*aac(3)IV+oriT*] resistance cassette. Figure 3.10 B shows that amplification of the desired 1020 bp product occurred in all putative mutant strains and further indicate that at

Figure 3.10 Verification of *sco3550* replacement in *S. coelicolor* M600. Chromosomal DNA was isolated from 7 putative mutants and subjected to PCR to confirm replacement of *sco3550* with the disruption cassette. (A) One microgram of chromosomal DNA and the primers KGI4 and KGI5 were used to amplify a 577 bp internal region of *sco3550* and products were resolved by 5% PAGE. *S. coelicolor* M600 served as a positive control and molecular weight marker V (Roche) was used as a size standard. (B) Primers JST5, annealing upstream of the predicted *sco3550* ATG start codon, and KGI6, annealing within the [*aac(3)IV+oriT*] disruption cassette, were used to generate a 1020 bp product that was resolved by 1.0 % agarose gel electrophoresis to test for replacement in the chromosome of the *sco3550* coding region by the cassette. Note that the mutagenized cosmid, pAU333, and the wild-type SCH5 cosmid were used as positive and negative controls, respectively while molecular wight marker III (Roche) was used as a size standard. (C) Crude extracts were isolated from putative *sco3550* mutant strain 20 and from M600 *S. coelicolor* after 24 hr growth on R2YE plate medium and 50 μg of each sample was separated by 6 % SDS-PAGE and examined by western analysis using affinity-purified anti-SCO3550 antibodies. A larger putative SCO3550 isoform (i) and a smaller SCO3550 isoform (ii) are indicated.

(A)


putative *sco3550* mutants
pAU333
MWMV
2
3
18
20
21
24
36
M600
577 bp

(B)


putative *sco3550* mutants
bp
MWMIII
SCH5
pAU333
2
3
18
20
21
24
36
1375
947
1020 bp

(C)

putative *sco3550* mutant 20
M600
(i)
(ii)

least a fraction of the genomic content of the putative mutants contained *aac(3)IV+oriT* disrupted *sco3550*. A failure to isolate strains in which all of the chromosomes contained *sco3550* disruptions was also substantiated by western analysis, which showed that the predicted native forms of SCO3550 were present in at least one and presumably all, of these strains (Figure 3.10 C).

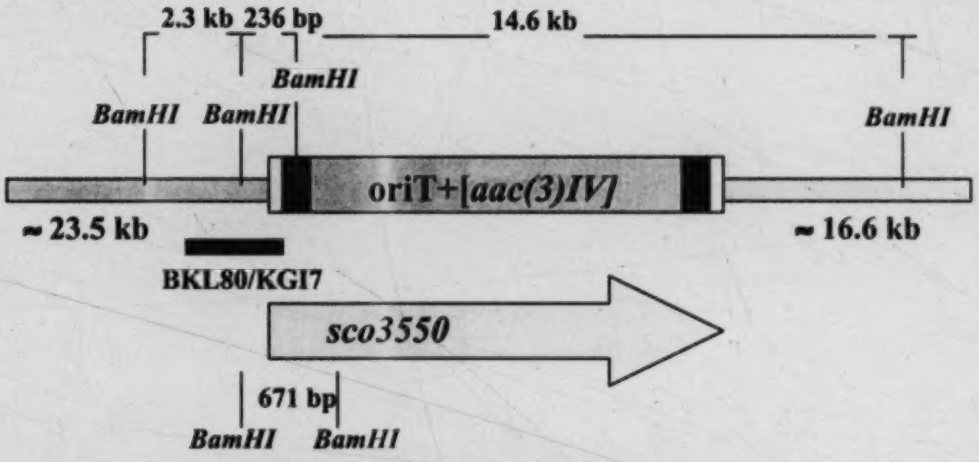
At this point, the inablility to isolate strains in which all chromosomes contained the desired *sco3550* disruption supported the earlier suggestion by Stoehr that *sco3550* has an essential function. However, to rule out a technical reason for the failure, the PCR-targeted mutagenesis protocol was again attempted but in *S. coelicolor* wild-type strain M145 rather than M600 under the presumption that the intrinsic differences between wild-type strains may facilitate *sco3550* disruption in M145. The *sco3550* disrupted SCH5 cosmid, pAU333, was transferred by conjugation into M145 spores and the exconjugants were again subjected to the regime of replica-plating in order to identify apramycin^R kanamycin^S colonies on DNA agar plates. As a result, several putative double cross-over exconjugants were isolated and spore suspensions were created followed by a subsequent round of replica-plating to isolate apramycin^R kanamycin^S clones arising from single spores. Interestingly, strain #2c exhibited a bald phenotype and strain #8a showed some delayed pigment production when plated on R2YE medium, in contrast to other potential mutants that continued to display a wild-type phenotype.

Verification of *sco3550* disruption was performed in apramycin^R kanamycin^S clones by Southern hybridization analysis and chromosomal PCR. For Southern analysis, chromosomal DNA was isolated from wild-type M145 and 8 putative mutant strains and digested with *Bam*HI. The digested DNA was then probed with a ^{32}P- random primer

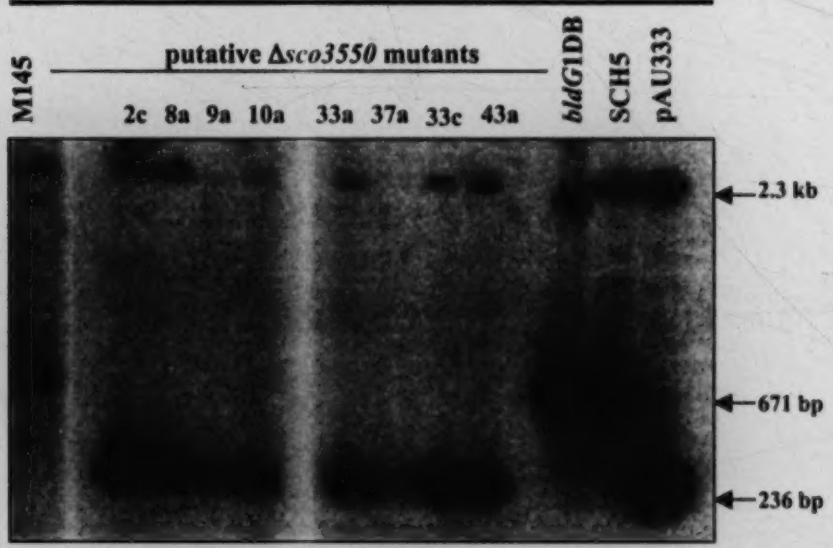
labeled probe generated through PCR amplification from a pAU227 template with primers KGI7 and BKL80 (Table 2.5). The probe corresponded to *bldG-sco3550* intergenic sequence and spanned approximately -320bp to +20 bp upstream of the predicted *sco3550* ATG start codon. As detailed in Figure 3.11 A, *sco3550* gene replacement in the chromosome was detected by the disappearance of a 671 bp *Bam*HI fragment and the appearance of a 236 bp *Bam*HI fragment. In contrast both wild-type and mutant strains retained a 2.3 kb fragment. Clearly, gene-specific hybridization analysis revealed *sco3550* was disrupted in at least 6 of the tested mutants (Figure 3.11 B). *Bam*HI digested chromosomal DNA from the M145 and mutant strains was also probed with the ^{32}P-labeled apramycin resistance cassette [*aac(3)IV+oriT*] and, as expected, revealed the presence of the cassette only in the mutant strains (Figure 3.11 C). This strongly suggested that apramycin resistance in the mutant strains was due to chromosomal disruption rather that spontaneous apramycin resistance. Although the Southern analysis seemed to suggest that all of the mutants contained the desired *sco3550* disruption and that wild-type copies of the gene were not present, the varying phenotypes of the mutant strains was difficult to explain. Chromosomal PCR was therefore performed in order to confirm the absence of wild-type copies of *sco3550* in chromosomal DNA preparations from the putative mutants. The chromosomal PCR was performed using primers hel-L and hel-R to amplify a 124 bp fragment internal to *sco3550* (Table 2.5)(Figure 3.12 A), and using primers JST5, which annealed upstream of the *sco3550* ATG start codon, and KGI6, which anneals within the apramycin [*aac(3)IV+oriT*] resistance cassette, to detect the presence of the integrated apramycin resistance cassette (Figure 3.12 B). As seen in Figure 3.12 A, a 124 bp product was

Figure 3.11 Verification of *sco3550* deletion in the M145 chromosome. (A) A schematic diagram representing the *Bam*HI digestion pattern for wild-type *sco3550* and that of the gene replaced by the [*aac(3)IV+oriT*] cassette [not to scale]. The location of the anealing site for the BKL80/KGI7 probe is indicated showing the overlap with the two chromosomal fragments liberated by *Bam*HI digestion. The *S. coelicolor* chromosomal sequence is indicated by the grey boxes. (B) and (C) show Southern analysis of the *sco3550* gene replacement. *Bam*HI-digested chromosomal DNA (2.5 μg) from 8 putative M145 mutant strains was separated by 1.0 % agarose gel electrophoresis. λ Pst was used as a molecular weight marker (not shown). After overnight transfer to a nylon membrane, the chromosomal DNA samples were probed with either ^{32}P- random primer labelled BKL80/KGI7 (B) or with labeled [*aac(3)IV+oriT*] cassette (C). The size and position of bands that were used to verify *sco3550* gene replacement are labelled and correspond to the *Bam*HI digestion pattern described in (A). *Bam*HI-digested samples of *bldG* 1DB chromosomal DNA and SCH5 cosmid DNA were used as negative controls for *sco3550* gene replacement and similarly digested M145 and pAU333 DNA samples were used as positive controls.

(A)


2.3 kb
236 bp
14.6 kb
BamHI
BamHI
BamHI
BamHI
oriT+[aac(3)IV]
≈ 23.5 kb
≈ 16.6 kb
BKL80/KGI7
sco3550
671 bp
BamHI
BamHI

(B)


32P-labelled BKL80/KGI7
M145
putative Δsco3550 mutants
2c
8a
9a
10a
33a
37a
33c
43a
bldG1DB
SCH5
pAU333
2.3 kb
671 bp
236 bp

(C)

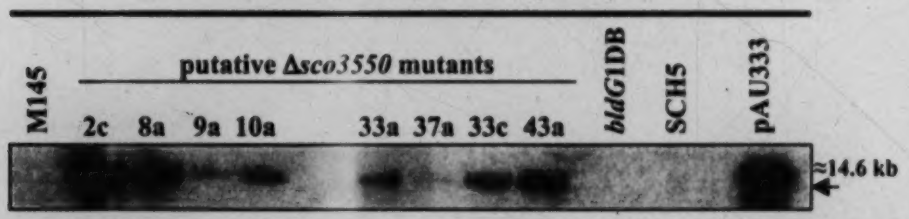
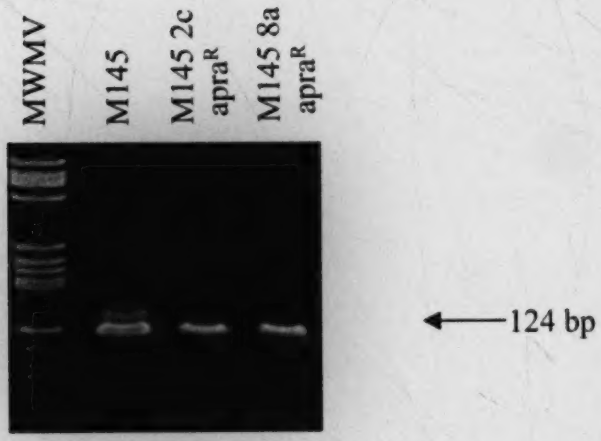
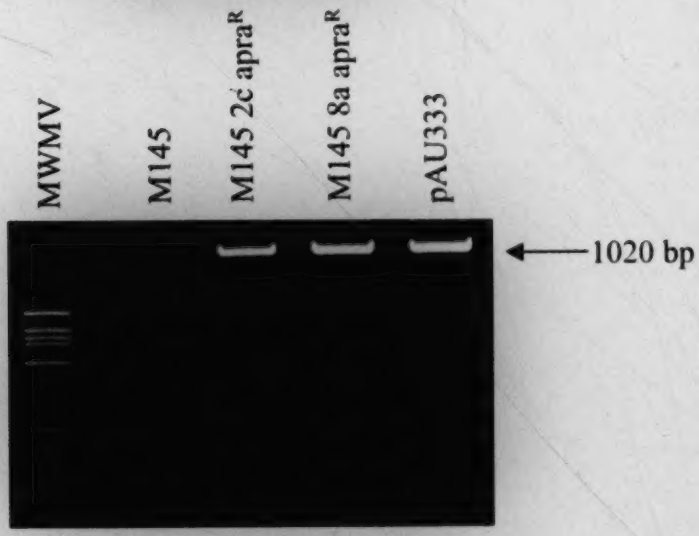

32P-labelled [aac(3)IV+oriT] cassette
M145
putative Δsco3550 mutants
2c
8a
9a
10a
33a
37a
33c
43a
bldG1DB
SCH5
pAU333
≈14.6 kb

Figure 3.12 Verification of *sco3550* replacement in *S. coelicolor* M145. Chromosomal DNA was isolated from putative mutant strains #2c and #8a and subjected to chromosomal DNA PCR. (A) Primers hel-L and hel-R (Table 2.4) were used to amplify a 124 bp internal region of *sco3550* and products were resolved by 5% PAGE. As a positive control PCR was performed using *S. coelicolor* M145 chromosomal DNA and Molecular weight marker V (MWMV) (Roche) was used as a size standard. (B) To test for the presence of the disruption cassette, primers JST5, annealing upstream of the predicted *sco3550* ATG start codon, and KGI6, annealing within the [*aac(3)IV+oriT*] disruption cassette, were used to generate a 1020 bp product that was resolved by 1.0 % agarose gel electrophoresis. The mutagenized cosmid, pAU333, and *S. coelicolor* M145 DNA were used as positive and negative controls, respectively while MWMV was used as a size standard. (C) Crude extracts were isolated from putative *sco3550* mutant strain #2c and #8a and *S. coelicolor* M145 after 24 hr growth on R2YE agar and 50 µg of each total cell protein sample was separated by 6 % SDS-PAGE and examined by western analysis using affinity-purified anti-SCO3550 antibodies. As in Figure 3.10, both the larger putative isoform (i) and smaller isoform of SCO3550 (ii) are as indicated.

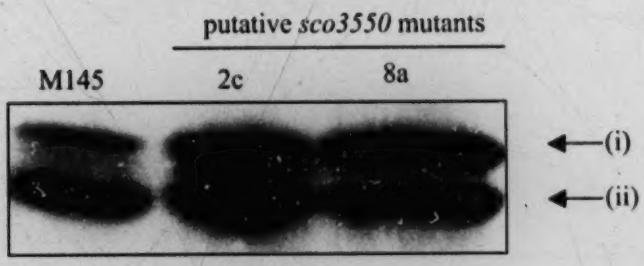
(A)


MWMV
M145
M145 2c apra^R
M145 8a apra^R
124 bp

(B)


MWMV
M145
M145 2c apra^R
M145 8a apra^R
pAU333
1020 bp

(C)


putative sco3550 mutants
M145
2c
8a
(i)
(ii)

observed which is representative of intact *sco3550*, while, Figure 3.12 B shows the expected 1020 bp apramycin resitance cassette disrupted *sco3550* in the putative mutant strains tested as well as in pAU333, indicating that both putative mutant strains appear to have a mixture of wild-type and *sco3550* disrupted chromosomes. In addition, western analysis again revealed the presence of the predicted native form of SCO3550 in all putative mutant strains tested (Figure 3.12 C).

The apparent absence of the wild-type *sco3550* gene as visualized by Southern analysis in contrast to the observed presence of the *sco3550* gene and protein by chromosomal PCR and western analysis, respectively, contends that a minor fraction of wild-type chromosomes still exist in the putative mutant strains. These wild-type chromosomes may be numerous enough to provide *sco3550* expression to complement the growing mycelium without being a large enough fraction of the total genomic complement to be detected by Southern analysis. Attempts were made to protect against this occurrence by the creation of stocks of unigenomic spores prior to molecular verification, however, it is possible that small mycelial fragments containing >1 *S. coelicolor* chromosomes may have circumvented mycelial filtration and led to the growth of supposed single spore isolates. Observations by Flardh (2003) reveal that for essential genes, such as *divIVA*, single cross-over integrants may arise that contain both the disrupted and wild-type version of the gene but have also been subjected to cosmid rearrangements and the loss of the kanamycin resistance cassette which yields a false interpretation of kanamycin sensitivity (Flardh, 2003). Although this may also be the case in the *sco3550* mutagenized strains, it must be occurring rarely as Southern analysis clearly displayed that the majority of chromosomal DNA tested contained a *sco3550*

disrupted ORF. In conclusion, the use of PCR-targeted mutagenesis was unable to completely disrupt *sco3550* in both *S. coelicolor* M145 and M600 strains and through mechanisms that are not entirely clear, at least some wild-type chromosomes were maintained in these strains.

While experiments were underway to verify the putative M145 *sco3550* disruption mutants, additional experiments were undertaken to try and complement the bald and delayed pigment production phenotypes of strains #2c and #8a. Attempted complementation experiments made use of recombinant plasmids containing a spectinomycin resistance gene (*aad*) to enable selection in mutant backgrounds containing the already integrated apramycin resistance cassette. Both pAU334 and pAU332 were constructed to allow exogenous expression of *sco3550* from the *attB* integration site in the chromosome of *S. coelicolor* (Table 2.4). For construction of pAU334, *sco3550* was liberated from pAU227 by *Sac*I/ *Nco*I digestion and later blunted to form a 3.2 kb fragment containing both 366 nt upstream of the predicted ATG start codon and 352 nt downstream of the stop codon. The blunted 3.2 kb fragment was then cloned into pSETΩ, which had been digested with *Eco*RI and later blunted. After verification of successful ligation by restriction enzyme digestion, the recombinant plasmid was designated pAU334 and transferred by conjugation into putative *sco3550* mutant strains #2c and #8a. pAU334 contained predicted *cis* regulatory regions of *sco3550* and it was hoped that the gene would be expressed at normal timing and levels albeit from a foreign location in the chromosome. Unfortunately, complementation was not observed and it was tentatively concluded that the strains may have contained second site mutations that created the observed phenotypes. This conclusion was later supported

by the finding that mutants #2c and #8a contained wild-type copies of the chromosome and should have displayed a wild-type phenotype. It was also a possibility that because the entire *bldG* locus was not included in the complemenation construct that *sco3550* might have been lacking wild-type regulation and not expressed at proper levels. To address this possibility, pAU332, which contains a thiostrepton inducible copy of *sco3550*, was constructed to enable control over *sco3550* expression. To construct pAU332, pAU331 was first constructed by moving the *sco3550* coding sequence into the pIJ6902 plasmid containing an apramycin resistance gene. It was necessary to clone the ATG start codon of *sco3550* in-frame with an *Nde*I site in pIJ6902 to ensure control by the thiostrepton-inducible p*tipA* promoter. A two part cloning strategy was performed similar to that used in construction of pAU330 (see Section 3.2, Figure 3.3). However, the initial PCR amplification of the 5' *sco3550* coding sequence used primer JST12-2, which included an engineered *Nde*I site at the ATG start condon, and primer BKL82. The resulting fragment was digested with *Nde*I and *Bam*HI and cloned into the similarly digested pIJ6902 vector and the cloning of the 3' *sco3550* coding sequence and downstream untranslated region proceeded as for pAU330. Replacement of the apramycin resistance gene with a spectinomycin resistance gene to construct pAU332 consisted of replacing the *Sac*I digested apramycin resistance cassette fragment with a *Sac*I digested *aad* fragment provided by D. Bignell. pAU332 was then transferred by conjugation from *E. coli* into mutant strains #2c and #8a. As a control, insert-free pIJ6902Ω [pIJ6902 with the apramycin resistance cassette replaced by the *aad* gene (Bignell et al., 2003)] was separately introduced into the mutant strains. The strains were grown on varying concentrations, ranging from 0-100 μg/mL, of thiostrepton inducer on

R2YE agar containing spectinomycin in order to induce *sco3550* expression, however, complementation was again not observed. Further examination of complemenation was terminated when the chromosomal PCR and western analyses suggested the presence of an intact *sco3550* gene in the putative mutant strains.

3.4.2 Construction of a conditional *sco3550* deletion mutant in *S. coelicolor*

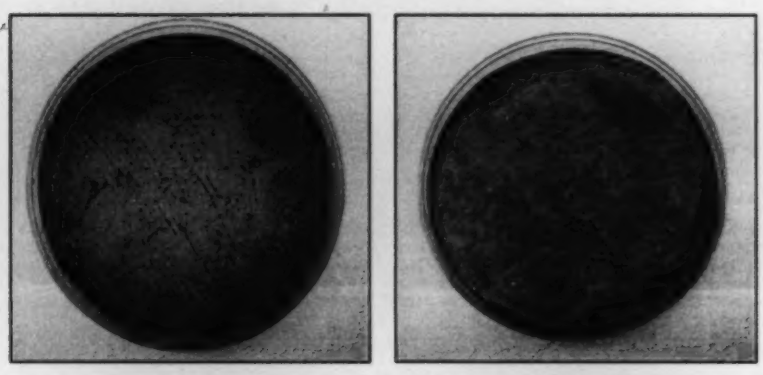
To provide definitive evidence that *sco3550* is an essential gene in *S. coelicolor*, conditional mutagenesis was attempted in order to disrupt the wild-type gene in the presence of a second inducible copy of *sco3550*. This was followed by an examination of the phenotypic effects observed in the disrupted strain when the extra copy of *sco3550* remained uninduced.

Conditional mutagenesis involved PCR-targeted mutagenesis of the chromosomal copy of *sco3550* in *S. coelicolor* M600 containing an inducible, extra copy of *sco3550*. In this experiment, pAU332, which contained the *sco3550* coding region downstream of the thiostrepton inducible p*tipA* promoter that had been introduced on an *att*-site integrating plasmid vector, and an insert-less control plasmid, pIJ6902Ω, were introduced by conjugation from *E. coli* into wild-type *S. coelicolor* M600.

Chromosomal integration of pAU332 was verified by chromosomal PCR (not shown). Prior to disruption of the native chromosomal copy of *sco3550* in the strain containing pAU332, the phenotypic effect of induction of an extra copy of *sco3550* in wild-type *S. coelicolor* M600 was examined on R2YE agar containing spectinomycin (to select for integrated plasmid maintenance) and a range of concentrations (0-100 μg/mL) of the thiostrepton inducer (Figure 3.13). The only visible change in phenotype induced

Figure 3.13 Phenotypic effect of SCO3550 overexpression in *S. coelicolor*. *S. coelicolor* M600 strains containing pAU332 or the control plasmid pIJ6902Ω were grown on R2YE agar containing varying concentrations (0-100 μg/mL) of thiostrepton inducer for approximately 48 hours.

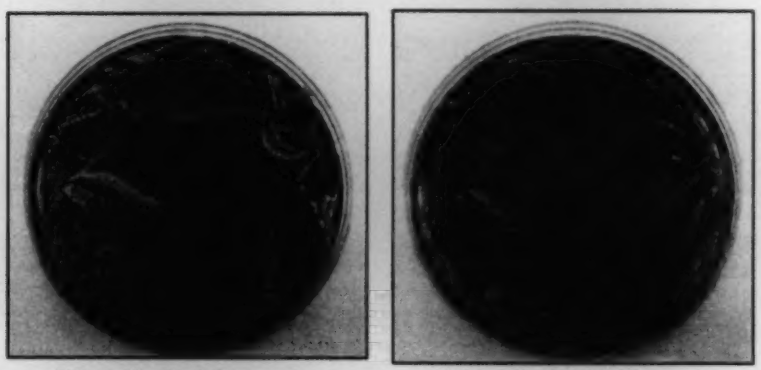
M600 + pIJ6902Ω



0 μg/mL thiostrepton 100 μg/mL thiostrepton

M600 + pAU332

0 μg/mL thiostrepton 10 μg/mL thiostrepton



30 μg/mL thiostrepton 100 μg/mL thiostrepton

by the presence of an extra copy of *sco3550* was a slight delay in development in *S. coelicolor*. When the pAU332-containing strains were grown for 2-3 days in the presence of increasing concentrations of thiostrepton they displayed considerably less aerial hyphal production, sporulation and blue pigment production relative to either the wild-type strain (not shown) or a strain containing the pIJ6902Ω control plasmid.

In order to generate the conditional mutant, cosmid SCH5 containing the apramycin resistance cassette [*aac(3)IV+oriT*] in place of the *sco3550* ORF (see Section 3.5.1) was transferred by conjugation into *S. coelicolor* M600 containing pAU332. Trans-conjugants were initially plated on 30 μg/mL thiostrepton to induce the second copy of *sco3550* prior to the addition after 16 hours of nalidixic acid and apramycin to select for transconjugants containing the *sco3550*-disruption. Initial thiostrepton induction of the extra copy of *sco3550* was carried out because it was believed that a higher proportion of wild-type chromosomes would be targeted by double cross-over replacement of *sco3550* with [*aac(3)IV+oriT*] if there was a background level of *sco3550* expression from the inducible copy of the gene. The expectation was that it should have been possible to isolate a strain with the native *sco3550* disrupted in all copies of the chromosome if a basal level of *sco3550* expression was maintained and that such a strain could be subsequently used to address the phenotypic effects of SCO3550 depletion by removal of the inducer of the *att*-site integrated copy of the gene. Ex-conjugants were restreaked on MS agar containing apramycin, nalidixic acid and 30 μg/mL thiostrepton and then replica-plated to DNA agar plates containing apramycin and thiostrepton with and without the addition of kanamycin. As well, the exconjugants were also restreaked onto a second set of agar plates containing nalidixic acid and apramycin

with and without kanamycin and all lacking thiostrepton. Conditional mutants arising from double-cross over events that brought about replacement of *sco3550* with [*aac(3)IV+oriT*] were expected to be apramycin^R kanamycin^S and to display a wild-type phenotype in the presence of thiostrepton. If *sco3550* played a role in morphological differentiation or antibiotic production those same mutants were expected to exhibit an aberrant phenotype when plated on medium lacking the thiostrepton inducer. However, repeated rounds of plating for single colony isolation and DNA agar replica-plating as well as further phenotypic analysis on R2YE plates with and without thiostrepton induction, failed to produce any isolates with altered phenotypes. Although it is possible that a *sco3550* disruption mutant was isolated that contained a disruption in all chromosomes and did not convey a visible phenotype, it is our belief that conditional mutagenesis using the pIJ6902 vector in this strategy exemplifies several technical limitations that will hinder the identification (if not the construction) of a putative conditional mutant (see Discussion).

3.4.3 Analysis of the *bldG*3b/ *sco3550* double mutant

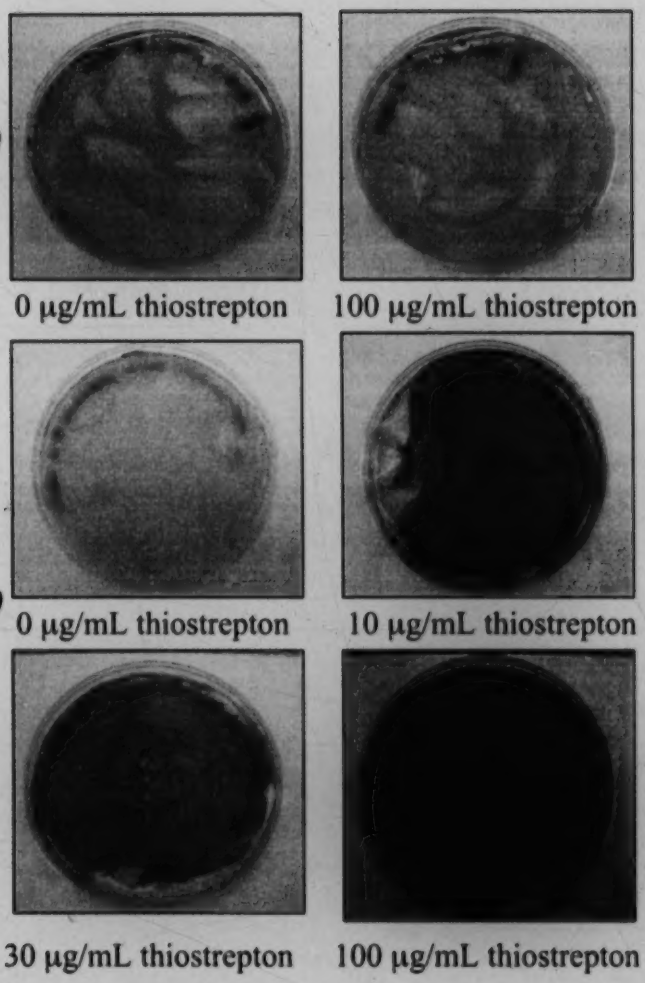
The supposition that *sco3550* is an essential gene and resistant to mutagenesis in studies by Stoehr was supported by the relative ease that *sco3550* could be disrupted in the *bldG*3b mutant background (Bignell et al., 2000; Stoehr, 2001). As already mentioned, it was postulated that SCO3550 may play a role in repressing genes involved in sporulation; therefore, in potential *sco3550* single mutant strains, sporulation is predicted to occur too early in the life-cycle causing lethality. Hypothetically, the double mutant is viable because the background disruption of *bldG* impedes the initiation of

development rendering a *sco3550* (repressor of sporulation) mutation inconsequential (Stoehr, 2001). It was suggested that the *bldG*3b/*sco3550* double mutant strain may be utilized in this study to construct a conditional *sco3550* single mutant strain.

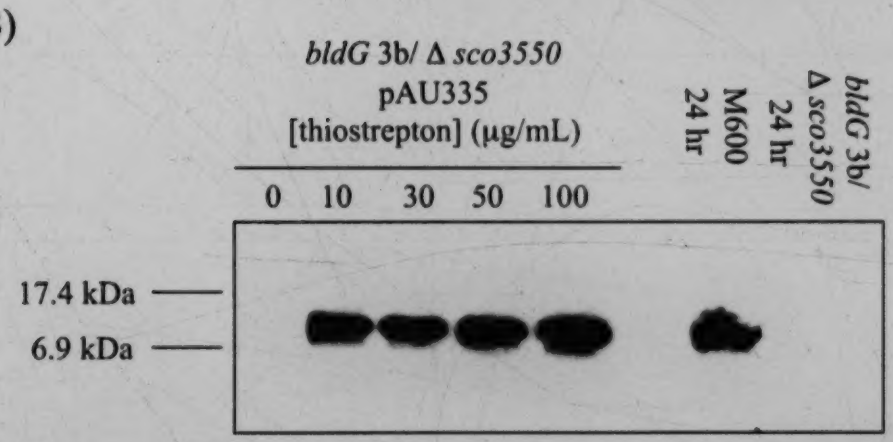
Our conditional mutagenesis strategy predicted that if disruption mutagenesis of *sco3550* was only possible in a *bldG* mutant background, then exogenous addition of BldG to the *bldG*3b/*sco3550* double mutant strain would effectively create a *sco3550* single mutant. The conditional *sco3550* single mutant could then assume three possible fates. Firstly, if a *sco3550* disruption strain is viable, the strain would be complemented by *bldG* expression and undergo a proper developmental program or secondly, even when *bldG* was added it would assume a non-wild-type phenotype that could be attributed to the loss of *sco3550*. If the gene is essential, the strain would most likely appear very sick or not be able to grow at all in the presence of exogenous BldG. The *bldG* inducible plasmid, pAU335, was constructed by ligation of a *NdeI*/ *Xba*I digested *bldG* fragment with the similarly digested pIJ6902Ω vector (Table 2.4). pIJ6902Ω was used in this experiment in order to impart spectinomycin resistance and allow selection of pAU335 in the double mutant as both thiostrepton and apramycin resistance genes had been used in the construction of the *bldG*3b/*sco3550* double mutant strain. pAU335 was transferred by conjugation from *E. coli* into the *bldG*3b/*sco3550* strain and exconjugants were restreaked onto R2YE agar containing spectinomycin and varying concentrations of thiostrepton inducer (ranging from 0-100 μg/mL) in order to induce expression of exogenous *bldG* from the p*tipA* promoter. As shown in Figure 3.14 A, partial complementation of the bald phenotype occured at higher levels of thiostrepton addition,

Figure 3.14 Conditional mutagenesis of *sco3550* in a *bldG*3b/*sco3550* double mutant background. (A) The *bldG*3b/*sco3550* double mutant harboring pAU335, which contains the *bldG* coding sequence under control of the p*tipA* promoter was grown on R2YE medium containing spectinomycin and varying concentrations of thiostrepton inducer, ranging from 0-100 μg/mL. (B) Crude extracts were isolated from the double mutant strain grown on varying concentrations of thiostrepton and western analysis was performed using anti-BldG antibodies. Samples (40 μg total protein) were separated by 15 % SDS-PAGE, transferred to a PVDF membrane and probed with OG1 anti-BldG polyclonal antibodies at a dilution of 1: 20 000. Subsequently, the membrane was incubated with Horseradish peroxidase-conjugated secondary antibody at a 1: 10 000 dilution and detection was performed by a 1 minute incubation with Enhanced plus chemiluminescence reagents and film exposure for 1 minute. Crude extracts isolated from a M600 wild-type strain and *bldG*3b/*sco3550* double mutant strain grown for 24 hr on R2YE agar were used as positive and negative controls, respectively, for BldG expression.

(A)


bldG 3b/ Δ *sco3550*
pIJ6902Ω
0 μg/mL thiostrepton
100 μg/mL thiostrepton
bldG 3b/ Δ *sco3550*
pAU335
0 μg/mL thiostrepton
10 μg/mL thiostrepton
30 μg/mL thiostrepton
100 μg/mL thiostrepton

(B)

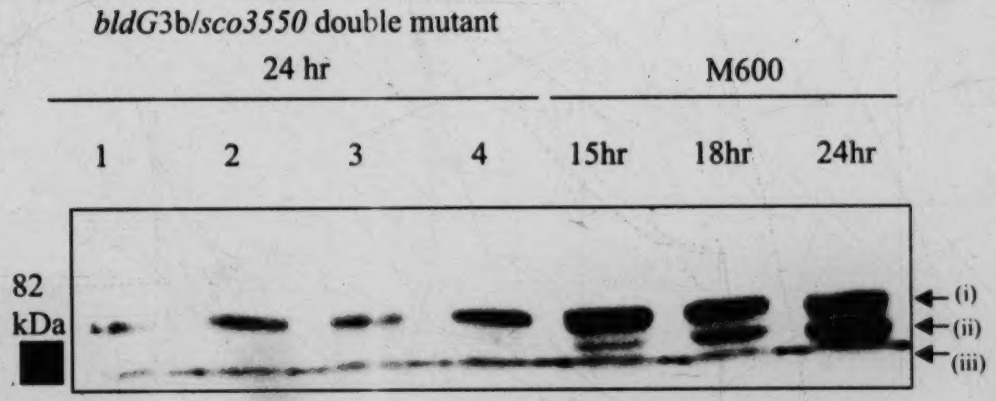

bldG 3b/ Δ *sco3550*
pAU335
[thiostrepton] (μg/mL)
0 10 30 50 100
M600
24 hr
bldG 3b/
Δ *sco3550*
24 hr
17.4 kDa
6.9 kDa

suggesting that the conditional *sco3550* mutant was viable. However, it should be noted that the presence of the inducer did not restore wild-type levels of pigmented antibiotic production and morphological differentiation to the bald *bldG*3b/*sco3550* double mutant. To confirm that complementation of the developmental delay in the putative double mutant strain was due to the induction of BldG, cell-free extracts were isolated from the strain grown at various concentrations of thiostrepton and western analysis was performed using anti-BldG antibodies. As shown in Figure 3.14 B, the presence of BldG protein was clearly induced by the addition of thiostrepton and induction increased slightly from 10 μg/mL to 50 μg/mL thiostrepton.

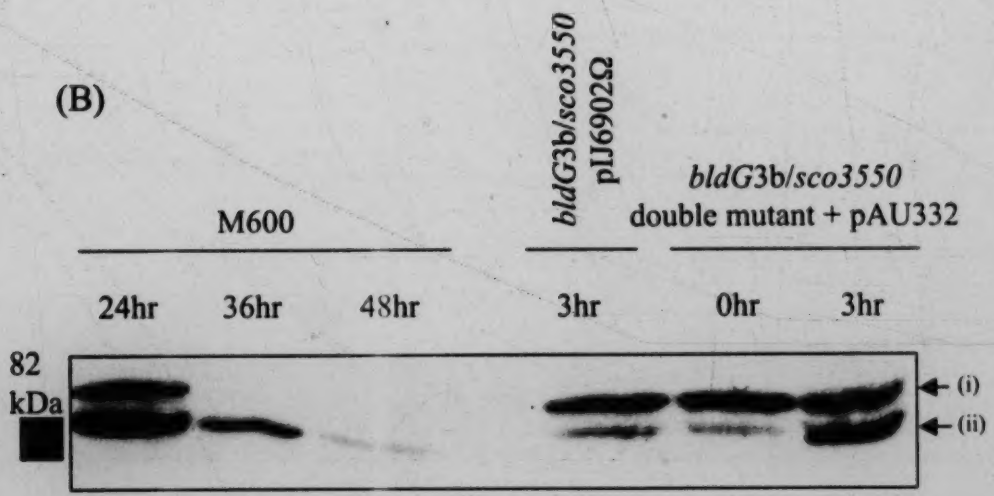
The above results suggest that a *sco3550* single mutation may in fact be viable however, in view of the previous failures to isolate a viable *sco3550* mutant, this finding called into question whether or not the *bldG3b/sco3550* double mutant was truly lacking in SCO3550 expression. Originally, *sco3550* gene disruption in the *bldG3b* mutant background had been confirmed by Southern analysis (Stoehr, 2001), however this study has shown that Southern analysis may be misleading and not representative of the entire chromosomal complement of the cell (see Figure 3.11 and 3.12). Therefore, the presence of SCO3550 in the *bldG*3b/*sco3550* double mutant was examined by western analysis. Twenty-four hour cell extracts from four separate protoplasted strains of the *bldG*3b/*sco3550* double mutant strain were examined for the presence of SCO3550. Protoplasting was performed in an effort to ensure that strains arose from a singe genome progenitor and could not contain a mixture of wild-type and disrupted chromosomes. Figure 3.15 A shows that all 4 strains examined by western analysis using affinity

Figure 3.15 Western analysis of cell extracts from *S. coelicolor bldG*3b/*sco3550* double mutant strains. (A) Crude extracts were isolated from four separate protoplasted strains of the *S. coelicolor bldG*3b/*sco3550* double mutant strain grown for 24 hr on R2YE agar. (B) (Figure 3.15B is reproduced from Figure 3.6 B and shows only part of the entire time course) The double mutant strain, which harbors the *sco3550* overexpression plasmid pAU332, was used to isolate crude extracts after a 0 hr and 3 hr thiostrepton induction of *sco3550* in R2YE liquid medium. Fifty micrograms of total cell protein were separated by 6% SDS-PAGE and then used in western analysis with affinity-purified anti-SCO3550 antibodies. The proteins were transferred to a PVDF membrane and probed with at least 10 µL of affinity-purified anti-SCO3550, and then by the addition of Horseradish peroxidase-conjugated secondary antibody at a dilution of 1: 10 000. Detection of hybridized bands was performed by exposure of the membrane to a 1 minute incubation with Enhanced plus chemiluminescence reagents and film exposure for 1 minute. Note that Kaleidoscope pre-stained standards were used as a size marker for SDS-PAGE and western hybridization. In panel (A), 15 hr-24 hr wild-type *S. coelicolor* M600 extracts were used as a positive control and in (B), 24 hr-48 hr wild-type *S. coelicolor* M600 extracts were used for the same purpose. The double mutant strain containing insert-less pIJ6902Ω total cell protein isolated after a 3hr thiostrepton induction was also used as a negative control for *sco3550* overexpression. (i) indicates the larger band of the doublet which is resolved by 6 % SDS-PAGE, (ii) the smaller band of the doublet and, at this point, presumed to be the SCO3550 specific reacting band from these experiments and (iii) indicates the dye front of the 6 % SDS-PAGE.

(A)


*bldG*3b/*sco3550* double mutant
24 hr
M600
1
2
3
4
15hr
18hr
24hr
82
kDa
(i)
(ii)
(iii)

(B)


*bldG*3b/*sco3550* pIJ6902Ω
*bldG*3b/*sco3550* double mutant + pAU332
M600
24hr
36hr
48hr
3hr
0hr
3hr
82
kDa
(i)
(ii)

purified anti-MBP-SCO3550 antibodies lack the lower band of the doublet that was observed by western analysis of wild-type crude extracts (see Section 3.3.2). These data suggests that the lower band of the doublet is in fact the native form of SCO3550 and that the double mutant strain indeed harbors a complete disruption of *sco3550*.

With these findings in mind, a parallel course of experimentation to identify a positive western control was performed by attempting to induce the expression of exogenous SCO3550 protein in the *bldG*3b/SCO3550 double mutant background. In this manner, pAU332, the pIJ6902 vector containing a spectinomycin resistance gene and *sco3550* under the control of the p*tipA* promoter, was used to analyze SCO3550 expression (Table 2.4). pAU332 was transferred to the double mutant strain by conjugation and exconjugants in which pAU332 was integrated into the *att*-site were selected on MS medium overlayed with 1 mg/mL spectinomycin. Chromosomal integration was confirmed by chromosomal PCR as described above (not shown). Positive strains were streaked for single colony isolates and liquid culture overexpression experiments were performed. The double mutant strain containing pAU332 was grown overnight in 50 mL R2YE broth and induced at a high OD with the addition of 30 μg/mL thiostrepton final concentration. Cell extract was harvested immediately after the addition of inducer as well as after 3 hours of induction. A strain containing the insert-free pIJ6902Ω vector was also harvested after a 3 hour induction to be used as a negative control in western analysis.

Western analysis of uninduced and induced cell extracts from the double mutant containing pAU332 revealed that only the lower band of the previously observed doublet was increased in the induced sample (Figure 3.15 B). This suggested that the lower band

in the resolved doublet is in fact the SCO3550 hybridizing band and that the upper band might represent cross-reaction with the antibodies. If true, the two negative controls, the double mutant strain containing pIJ6902Ω alone and the uninduced test strain, would be expected to lack the lower band. However, as seen in Figure 3.15 B both control strains exhibited a faint hybridizing band corresponding to the position of the lower band in the doublet. To try and reconcile the differing results seen in Figure 3.15 A and 3.15 B the experiments were repeated and it was found that in long film exposures that the lower band of the doublet could be observed in all putative *bldG*3b/*sco3550* double mutant strains albeit in varying strain-dependent amounts. Therefore, the conclusion that was reached is that the *bldG*3b/*sco3550* double mutant isolated by Stoehr is not a true *sco3550* null mutant and that it must house a small percentage of wild-type (with respect to *sco3550*) genomes due to the multiple genomic nature of *S. coelicolor* (see Section 3.5 mutagenesis). How this could be true after protoplasting, which is reported to result in the isolation of single genome-containing protoplasts, is difficult to explain. It is possible that the disruption is highly unstable and can flip out of the chromosome and regenerate the *sco3550* ORF by successive rounds of replication. This instability may account for the discrepancies seen by western analysis in the level of SCO3550 between double mutant strains that have been recently protoplasted versus strains that have been repeatedly cultured.

No discrete conclusions can be drawn from the conditional mutagenesis experiments regarding the possible essential nature of *sco3550*. The possibility exists that the strain contains a minor fraction of wild-type chromosomes that could not be detected by Southern analysis but which still resulted in the production of sufficient

amounts of the putative helicase protein to maintain viability of the strain. Stability of the *bldG*3b/*sco3550* double mutant strain is also an issue as the *sco3550* disruption was generated by a single cross-over between a internal gene fragment present in a non-replicating thiostrepton resistance gene containing plasmid and the native *sco3550* gene in the *S. coelicolor* chromosome and therefore the wild-type gene may be regenerated if the internal fragment and resistance gene flips out of the chromosome. This instability may account for varying levels of SCO3550 expression that have been observed between separate strains derived from a single parent strain. Furthermore, the background level of SCO3550 protein present in the supposed double mutant was examined by western analysis over the course of the BldG induction platings and was observed to be constant regardless of the level of thiostrepton induction (data not shown). Therefore, the stability of the SCO3550 single cross-over disruption does not appear to be affected by exogenous *bldG* expression and may suggest that there is sufficient SCO3550 protein present to perform its cellular role.

3.5 Western analysis revisited

3.5.1 Western analysis with a positive control

Although there appears to be a background level of SCO3550 present in the *bldG*3b/*sco3550* double mutant, crude extracts isolated from a 3 hour induction of the double mutant containing the pAU332 *sco3550* overexpression vector can still be used as a positive control for western analysis. In contrast to the control strains, induction of *sco3550* expression in the pAU332 containing strain is clearly seen as an increase in the abundance of the lower band of the observed doublet (see Section 3.4.3). Using the strain

containing pAU332 as a positive control, expression of SCO3550 was re-examined (Figure 3.16), and found to be similar to the previously observed western analysis using non-affinity purified antibodies (See Figure 3.5). Differences between the pattern observed in Figure 3.5 and that in Figure 3.16 could largely be explained by the resolution of the helicase into two bands using 6% SDS-PAGE. SCO3550 expression occurs throughout growth and development on R2YE agar, gradually increasing to peak protein levels at the 24 hour time point and then recedes through the latter portion of the life cycle. It is also a possiblity that the larger hybridizing band is representative of a larger isoform of SCO3550 and that *sco3550* is expressed as two separate forms.

Support for arguing the presence of two isoforms of SCO3550 comes from the presumed location of the SCO3550 translational start site. As aforementioned, the predicted translational start site of SCO3550 was shifted downstream 114 nt from the *S. coelicolor* annotated genome database predicted GTG translational start site to an in-frame ATG codon (Figure 3.17). The primary reason for this change in start site positioning arose from primer walking RT-PCR transcriptional start site localization experiments by Stoehr (Stoehr, 2001). These experiments used a variety of 5' oligonucleotide primers that were increasingly shifted towards the presumptive transcript start region and a single static 3' primer (JST1) that annealed within the *sco3550* transcript. The principle being that a 5' primer that annealed to *sco3550* transcript would produce a radiolabelled PCR product, whereas, once the primers were extended past the 5' transcription start site a product would not be amplified by RT-PCR. In this manner the transcription start site was localized to a 20 nt window flanked by primers JST5 and JST6 including the region of DNA containing the genome project predicted GTG

Figure 3.16 Western analysis revisited. This figure was generated by reproduction of Figure 3.6 B and 3.15 B to bring together western analysis of the SCO3550 protein expression profile with a positive control. Crude extracts isolated from growth on R2YE solid medium after various time-points corresponding to vegetation growth and differention in wild-type *S. coelicolor* M600 were separated by 6 % SDS-PAGE and used for western analysis with affinity-purified anti-SCO3550 antibodies. Crude extracts isolated from a 3 hr thiostrepton induction of the *S. coelicolor* strain containing the insert-free pIJ6902Ω and the *sco3550* overexpressing strain (*bldG*3b/*sco3550* + pAU332) at 0 hr induction were used as negative controls. Crude extract isolated from *bldG*3b/*sco3550* + pAU332 after a 3 hr induction of SCO3550 expression in liquid culture served as a positive control. The larger band in the doublet (i) and the induced lower band (ii) are indicated for the positive control.

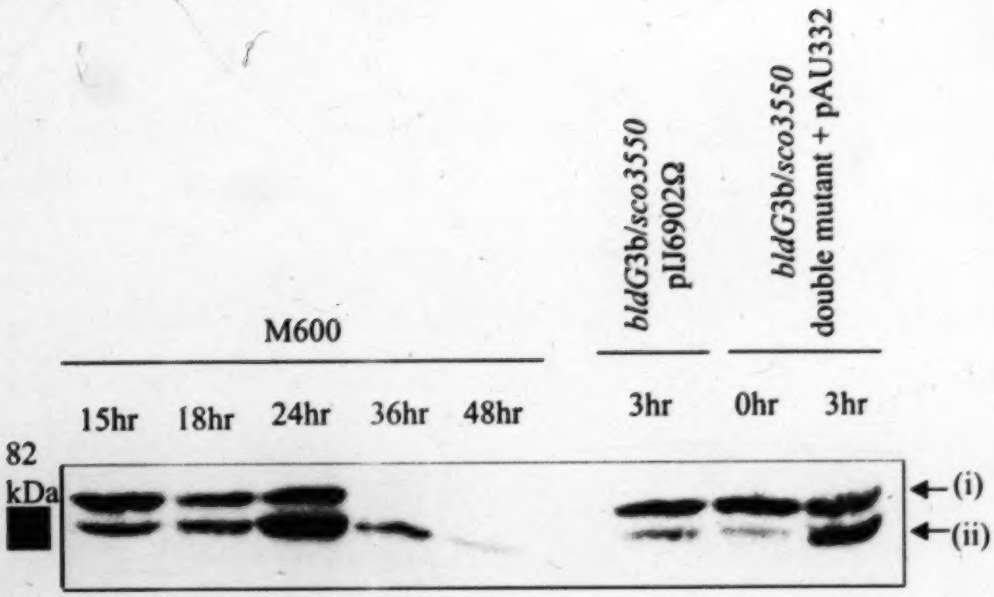
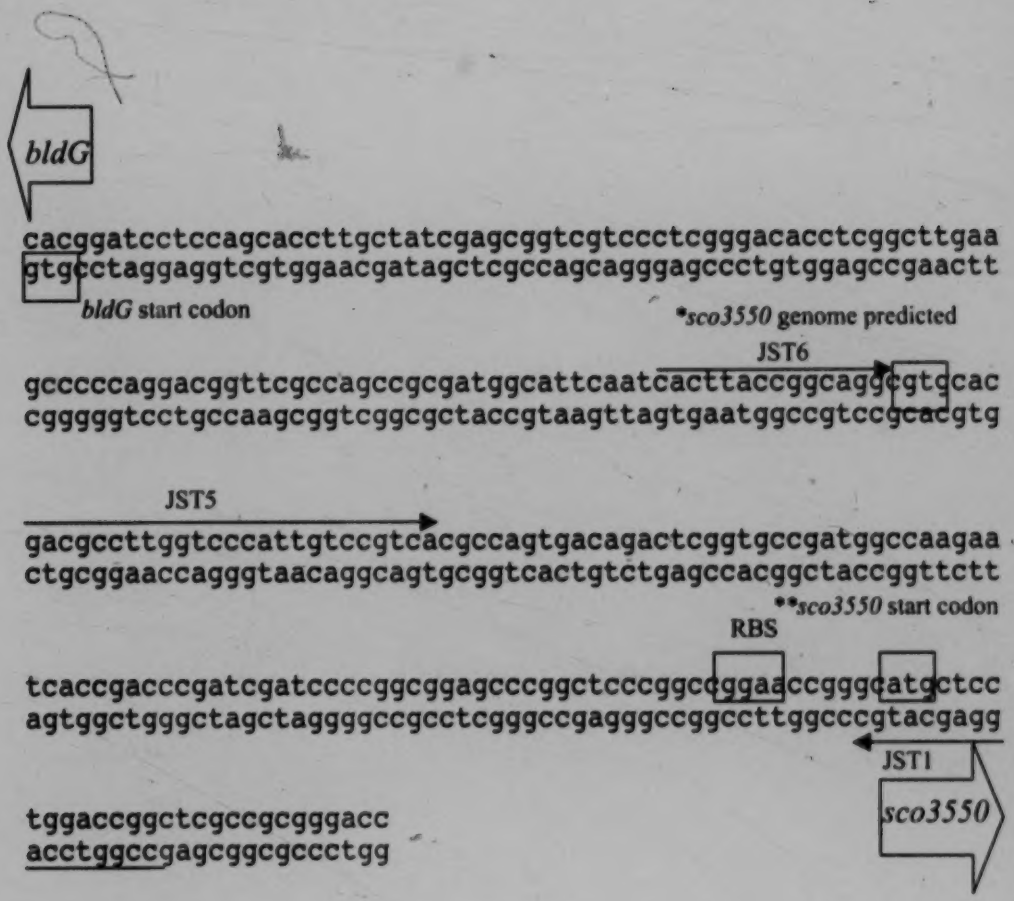

M600
bldG3b/sco3550 pIJ6902Ω
bldG3b/sco3550 double mutant + pAU332
15hr
18hr
24hr
36hr
48hr
3hr
0hr
3hr
82 kDa
(i)
(ii)

Figure 3.17 DNA sequence of the *bldG- sco3550* intergenic region. The double-stranded DNA sequence of the *bldG- sco3550* intergenic region is displayed. The *bldG* start codon, the GTG *sco3550* start codon predicted by the *S. coelicolor* genome sequencing project (http://www.sanger.ac.uk/Projects/S_coelicolor/) and the designated ATG *sco3550* start codon 114 nt downstream are labelled. Also, shown is a suitable purine-rich Shine-Dalgarno ribosome binding site (RBS) slightly upstream of the putative ATG start codon. Primers used by J. Stoehr in her primer walking and transcriptional start site localization studies are also indicated by long arrows, which represent the direction of oligonucleotide polarity from 5'to 3' and the region of complementarity. JST1, the static 3' primer for RT-PCR is also shown.


bldG
cacggatcctccagcaccttgctatcgagcggtcgtccctcgggacacctcggcttgaa
gtgcctaggaggtcgtggaacgatagctcgccagcagggagccctgtggagccgaactt
bldG start codon
*sco3550 genome predicted
JST6
gcccccaggacggttcgccagccgcgatggcattcaatcacttaccggcaggcgtgcac
cgggggtcctgccaagcggtcggcgctaccgtaagttagtgaatggccgtccgcacgtg
JST5
gacgccttggtcccattgtccgtcacgccagtgacagactcggtgccgatggccaagaa
ctgcggaaccagggtaacaggcagtgcggtcactgtctgagccacggctaccggttctt
**sco3550 start codon
RBS
tcaccgacccgatcgatccccggcggagcccggctcccggccggaaccgggcatgctcc
agtggctgggctagctaggggccgcctcgggccgagggccggccttggcccgtacgagg
JST1
tggaccggctcgccgcgggacc
acctggccgagcggcgccctgg
sco3550

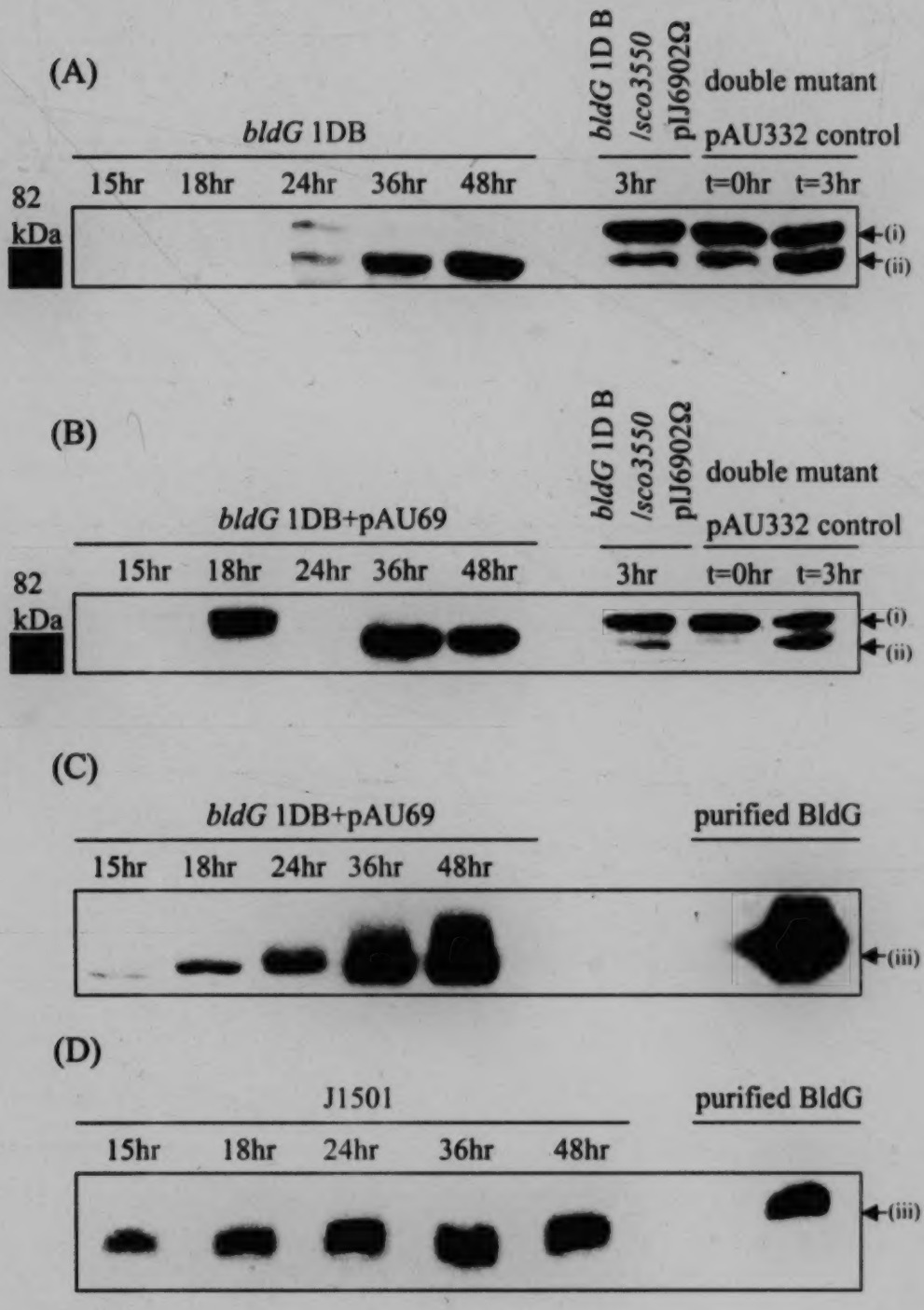
translational start codon. Due to this result, it was inferred that the predicted translational start site was simply too close to the localized transcriptional start site, although the rare exception of being a leaderless transcript was not investigated. Since leaderless transcripts are rare, the downstream ATG in-frame codon was designated the SCO3550 translational start site. This shift in start site was also supported by the presence of a purine-rich ribosome-binding site (RBS) upstream of the ATG codon, while, the predicted GTG codon did not have an accompanying predicted RBS. The recent visualization by western analysis of two anti-SCO3550 antibody-reacting bands calls into question the reassignment of the start codon. The expression of SCO3550 from the GTG predicted translational start site, as well as the designated ATG start codon, could account for the appearance of two separate isoforms of the proteins, with approximate calculated sizes of 87.69 and 86 kDa. In Figure 3.17, the putative larger isoform of the protein appears prevalent in the vegetative stages of growth and is completely absent in later time points, while, the smaller isoform increases to a peak level at 24 hours and then gradually recedes. Therefore, separate temporal profiles may imply the potential for separate activities or possible regulation of the smaller isoform by proteolytic processing of the larger species.

3.5.2 Western analysis of SCO3550 in *bldG* 1DB mutant extracts

Due to the close genomic proximity of *sco3550* to the *bldG* locus, it is a possibility that BldG and ORF3 may affect the expression of SCO3550 and changes to the SCO3550 expression may be observed in a *bldG* deletion mutant where there is no functional BldG protein present. To analyze the effects of a *bldG* 1DB in-frame deletion

on the expression profile of SCO3550, *bldG* mutant cell extracts were isolated over a 48 hour time course, and 50 μg of each sample were separated by 6 % SDS-PAGE, transferred to a PVDF membrane and hybridized with affinity-purified anti-SCO3550 polyclonal antibodies (Bignell et al., 2000). As seen in Figure 3.18 A, the expression profile of SCO3550 in the *bldG*1DB mutant was both greatly altered and delayed in comparison to the pattern previously observed in the wild-type strain (Figure 3.16). SCO3550 expression appeared markedly delayed in the *bldG* 1DB mutant being faintly detected at 24 hours and showing an increase in the amount of the lower putative isoform at 36 and 48 hours. This delay in protein expression is most likely not due to growth related effects as the *bldG* 1DB mutant strain customarily grows as fast as wild-type *S. coelicolor*, even though the mutant strain does not undergo morphological or physiological differentiation (personal communication, D. Bignell). An alternative interpretation of the SCO3550 expression profile in the *bldG* mutant extracts is that the profile has been reversed in comparison to that in the wild-type with SCO3550 being repressed in early growth and induced in later developmental time-points. Another striking observation, was the alteration in the balance or expression of the two putative isoforms of SCO3550. To recall the wild-type expression profile of SCO3550 (see Fig. 3.16), the larger putative isoform is present at a high level throughout vegetative time points and initial development, whereas, the smaller protein is gradually increased to peak levels at 24 hours and then notably recedes afterwards. However, in *bldG* 1DB mutant cell extracts, the larger putative isoform of SCO3550 is only very faintly present at 24 hours and absent from the rest of the time course; the smaller isoform has very low expression commencing at the 24 hour time point that increases to much higher levels at

Figure 3.18 Western analysis of SCO3550 in *bldG* 1DB mutant extracts. Crude extracts were isolated at varying time points over a 48 hour R2YE plate culture time course for the (A) in-frame *bldG* 1DB mutant strain or (B) *bldG* 1DB mutant strain complemented with pAU69 (encoding *bldG* expressed from its native promoter). The cell extract samples (50 μg total cell protein) were separated by electrophoresis on a 6 % SDS-polyacrylamide gel. After transfer to a PVDF membrane, the samples were probed with at least 10 μL of affinity-purified anti-SCO3550 antibodies, followed by a secondary hybridization with Horseradish peroxidase-conjugated secondary antibody at a dilution of 1: 10 000. Kaleidoscope pre-stained standards were used as size markers. In panel (A) and (B), pre and post-induction cell extracts from the *bldG* 3b/ SCO3550 double mutant containing the insert-free pIJ6902Ω or the inducible *sco3550* expression plasmid pAU332 were included as controls. The putative larger isoform of SCO3550 (i) and smaller isoform (ii) are indicated. Western analysis of BldG was also performed on 40 μg of crude extract isolated from strain *bldG* 1DB (complemented with *bldG* expressed from its own promoter on pAU69) (C) and from wild-type strain J1501 (D) [Figure reproduced with permission from (Bignell, et al., 2003)]. In this case the samples were subjected to 15 % SDS-PAGE, transferred to a PVDF membrane and probed with OG1 anti-BldG polyclonal antibodies at a dilution of 1: 20 000. Subsequently, the membrane was incubated with Horseradish peroxidase-conjugated secondary antibody at a 1: 10 000 dilution and detection was performed by a 1 minute incubation with Enhanced plus chemiluminescence reagents and film exposure for 1 minute. Kaleidoscope pre-stained standards were used as a size marker and in panel (C) and (D), FPLC-purified BldG (obtained from D. Bignell) was used a positive control and the BldG specific hybridizing band is indicated (iii).


(A)
bldG 1DB
bldG 1D B /sco3550 pIJ6902Ω
double mutant
pAU332 control
15hr 18hr 24hr 36hr 48hr
3hr t=0hr t=3hr
82 kDa
(i)
(ii)
(B)
bldG 1DB+pAU69
bldG 1D B /sco3550 pIJ6902Ω
double mutant
pAU332 control
15hr 18hr 24hr 36hr 48hr
3hr t=0hr t=3hr
82 kDa
(i)
(ii)
(C)
bldG 1DB+pAU69
purified BldG
15hr 18hr 24hr 36hr 48hr
(iii)
(D)
J1501
purified BldG
15hr 18hr 24hr 36hr 48hr
(iii)

36 and 48 hours. Specificity controls again included the previously mentioned *bldG*3b/*sco3550* double mutant with and without induction of *sco3550* expression from pAU332.

3.5.3 Western analysis of SCO3550 in *S. coelicolor bldG* 1DB + pAU69 complemented extracts

To confirm that SCO3550 protein expression was affected as a consequence of the disruption of *bldG* in the *bldG* 1DB mutant strain rather than due to growth related effects, western analysis of SCO3550 in extracts of *bldG* 1DB containing pAU69, which has *bldG* under control of its native promoter, was undertaken. The protein expression profile of SCO3550 was only partially restored, although not to wild-type levels (Figure 3.18 B), when the *bldG* mutation was complemented. Upon further examination it was observed that the larger putative isoform of SCO3550 showed low level expression at 15 hours that increased greatly in the18 hour time point. The presence of the smaller putative SCO3550 isoform still appeared confined to 36 and 48 hour time. It was suggested that the inability to restore the wild-type SCO3550 protein profile in the complemented *bldG* mutant strain was due to minor BldG expression differences between the wild-type and complemented *bldG* mutant. These proposed discrepancies in BldG expression might be manifested as a lack of complete complementation and delay in the restoration of morphogenesis in the complemented strain, although still maintaining similar growth levels as the wild-type strain (personal communication, D. Bignell). To test this possibility, western analysis was performed on cell extracts from *bldG* 1DB containing pAU69 to examine BldG expression relative to the wild-type

expression profile. Western analysis proceeded by separating 40 μg of complemented cell extract by 15 % SDS-PAGE, transfer to a PVDF membrane, and a subsequent primary hybridization with OG1 anti-BldG antibodies at a 1 in 20, 000 dilution. It was observed that the protein profile of BldG in the complemented extracts was delayed relative to the wild-type BldG expression course (Figure 3.18 C) [Figure 3.18 D (Bignell et al., 2003)], with little BldG expression observed at 15 and 18 hour time points. This difference between complemented and wild-type BldG expression profiles early in the life cycle, can most likely be attributed to the genomic differences between the location of the exogenous and wild-type *bldG* gene. In this respect, pAU69, although containing wild-type *bldG* under control of its endogenous promoter, integrates at the *attB* site and therefore, may lack *cis*-acting genomic determinants that influence wild-type BldG expression. This delay most likely accounts for the delay in complementation of *bldG* 1DB and possibly the partial restoration of the SCO3550 profile observed in pAU69-complemented extracts. It is also a possibility that the genomic separation of exogenous *bldG* and native *sco3550*, might be a factor in SCO3550 expression, if a coordinate control mechanism exists between the two genes. The results reported here support the belief that *bldG* imparts either a direct or indirect influence on the protein expression profile of SCO3550 in *S. coelicolor* and adds more support that there is a significant link between *sco3550* and the overlapping *bldG* locus.

3.6 Transcript analysis of *sco3550*

Analysis of *sco3550* gene expression was attempted to examine the timing of expression over development and accurately position the transcriptional start site. This

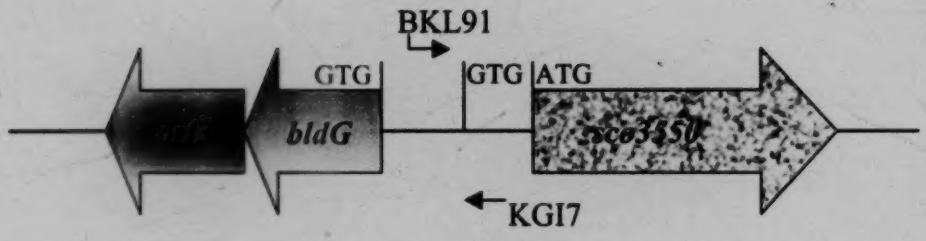
would include attempted verification of RT-PCR observations by Stoehr of a biphasic gene expression pattern displaying an initial 15 hour vegetative peak followed by a 36 hour peak occurring at a key transition point in development (Stoehr, 2001). In light of recent findings through western analysis, gene expression was also examined to further elucidate the possibility of multiple transcriptional start sites or transcript processing involved in *sco3550* expression.

3.6.1 S1 nuclease mapping of *sco3550* transcripts

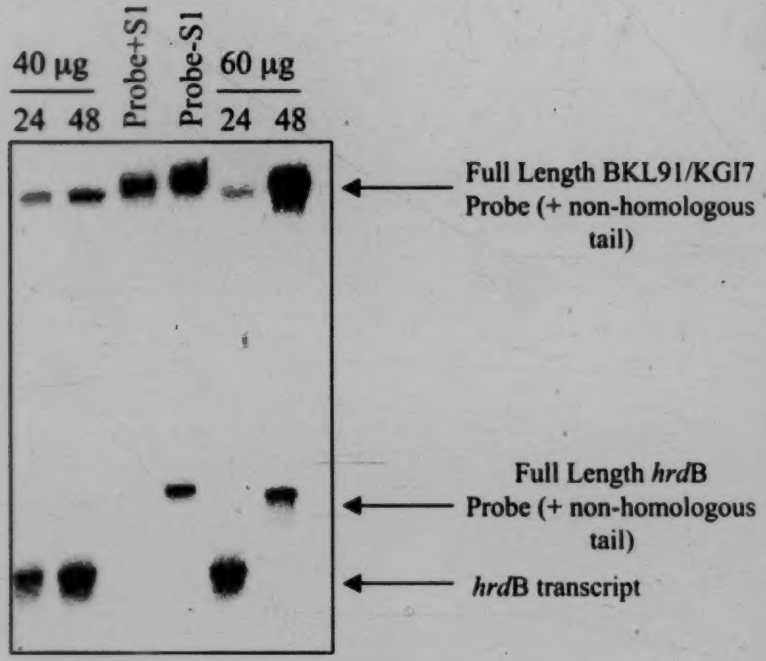
5' S1 nuclease mapping of *sco3550* transcripts was re-examined, initially with a probe that was constructed to include the predicted transcriptional start 'region' as localized through RT-PCR by Stoehr (Stoehr, 2001)(see Figure 3.17). The probe was an end-labelled 160 bp product generated by PCR amplification using pAU334 as template and BKL91 and KGI7 as primers. BKL91 contained a non-homologous 5' extension to discern between probe-probe reannealing and full length protection (Figure 3.19 A). The resulting probe extended from 32 bp upstream of the predicted *sco3550* ATG start codon to 40 bp upstream of the *bldG* start codon in the *bldG – sco3550* intergenic region. Hybridization was performed overnight at 45 °C with both 40 and 60 µg of total RNA, purified from 24 and 48 hour *S. coelicolor* M145 liquid R2YE cultures. These time points were chosen to decrease the number of samples required for preliminary mapping studies but to still examine both a vegetative and developmental stage of expression. An end-labeled *hrd*B specific probe designed to recognize the internal region of the *S. coelicolor* constitutively expressed vegetative sigma factor was also included in the hybridization reaction mixture to serve as an internal control for quantification, integrity

Figure 3.19 5' S1 mapping of *sco3550*. (A) Attempted 5' S1 nuclease mapping of SCO3550 transcripts utilized a 160 bp probe generated as a PCR product from pAU334 using primers BKL91 and KGI7. Primer positions relative to the predicted *sco3550* GTG and ATG start codon and *bldG* GTG start codon are indicated by small black arrows and the non-homologous extension on BKL91 is indicated as a vertical line attached to the arrow. (B) S1 mapping was performed on 40 μg or 60 μg of *S. coelicolor* M145 RNA isolated from a liquid culture time course at 24 and 48 hours post-inoculation and hybridization was performed overnight with end-labeled BKL91/ KGI7 probe at 45 °C. This was followed by S1 digestion of the resulting DNA/RNA hybrids and reactions were then precipitated and prepared for electrophoresis on a 6 % sequencing gel. Also included on the gel as controls for visualization of probe artifacts and probe-probe reannealing were samples that consisted of probe alone that had either been exposed to S1 nuclease digestion (Probe+S1) or not (Probe-S1). The locations of full length probe, representing probe-probe reannealing, for both the BKL91/KGI7 and the *hrdB* internal control probe are indicated as well as the detected *hrdB* transcript.

(A)


BKL91
GTG
GTG
ATG
bldG
sco3550
KGI7

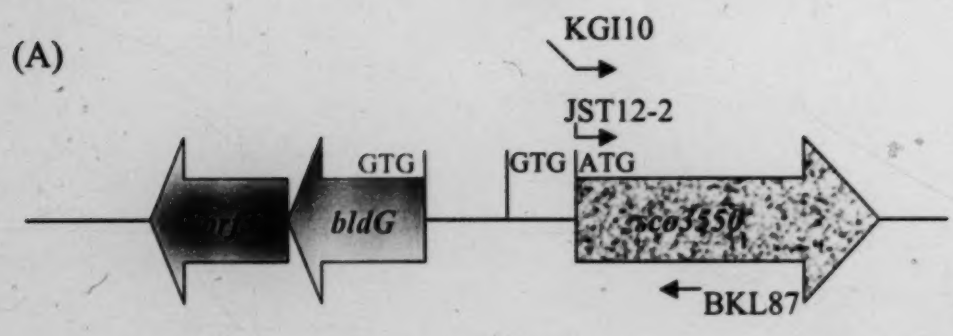
(B)

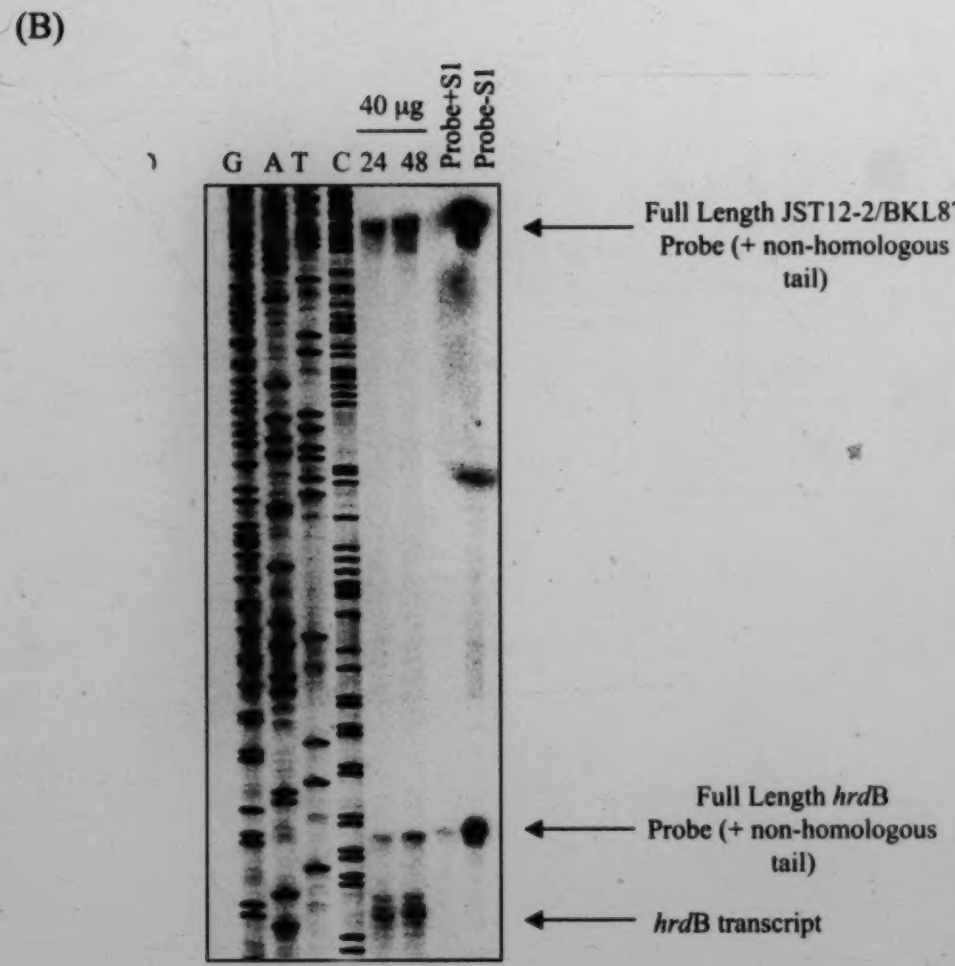

40 µg
60 µg
24
48
Probe+S1
Probe-S1
24
48
Full Length BKL91/KGI7 Probe (+ non-homologous tail)
Full Length hrdB Probe (+ non-homologous tail)
hrdB transcript

of RNA and efficacy of mapping. Upon completion of hybridization, samples were digested with S1 nuclease and separated by electrophoresis on a 6 % sequencing gel in conjunction with sequencing samples generated by manual sequencing with primer KGI7. Unfortunately, *sco3550* transcripts could not be detected by probe BKL91/ KGI7 protection of the analyzed RNA samples, as well as other plate culture samples tested, while the internal *hrdB* control transcript was clearly defined indicating that there were no wide-spread issues with RNA stability or quantification (Figure 3.19 B).

It was suggested that the inability to detect the 5' *sco3550* transcript might have been due to inherent low expression of the gene, extensive 5' RNA secondary structure or that the transcriptional start site may lay downstream of the designed KGI7-BKL91 probe. Consequently, S1 nuclease protection experiments were performed using an internal *sco3550* probe, just to detect transcripts and circumvent the issue of transcriptional start site prediction. Internal S1 nuclease protection analysis was performed using 40 μg of the same 24 and 48 hour RNA samples isolated from liquid culture of *S. coelicolor* M145 and used the end-labeled probe JST12-2/ BKL87 which started at the *sco3550* predicted ATG start codon and extended 224 bp into the *sco3550* ORF (Figure 3.20 A). As before, the probe contained a non-homologous extension on one end which is critical, especially for internal S1 nuclease protection analysis, to discriminate between full length probe protection of transcript and probe-probe reannealing. The *hrdB* internal control was also included in the reactions. S1 nuclease digested samples were prepared and separated by electrophoresis on a 6 % sequencing gel in conjunction with sequencing samples generated from manual sequencing with primer BKL87. Once again, the *sco3550* transcript could not be detected while the *hrdB*

Figure 3.20 Internal S1 nuclease protection analysis of *sco3550* transcripts. (A) The 224 bp probe was generated as a PCR product from pAU334 using primers JST12-2 and BKL87. Primer positions relative to the predicted *sco3550* GTG and ATG start codon are indicated by small black arrows, and the non-homologous extension on JST12-2 is indicated as a vertical line attached to the arrow. Primer KGI10 is also indicated as a black arrow with a larger non-homologous extension and was used to generate probe KGI10/ BKL87 that was used in a later round of internal S1 nuclease protection analysis. (B) S1 nuclease protection analysis was performed on 40 μg of *S. coelicolor* M145 RNA isolated from a liquid culture time course at 24 and 48 hours post-inoculation and hybridization was performed overnight with end-labeled JST12-2/ BKL87 probe at 45 °C. This was followed by S1 digestion of the resulting DNA/RNA hybrids and reactions were then precipitated and subjected to electrophoresis on a 6 % sequencing gel. Also included on the gel as controls for visualization of probe artifacts and probe-probe reannealing were samples that consisted of probe alone that had either been exposed to S1 nuclease digestion (Probe+S1) or not (Probe-S1). Manual sequencing reactions (G,A,T,C) generated using primer BKL87 were also included to identify accurately full-length protection and any detectable transcript present. The locations of full length probe, representing probe-probe reannealing, for both JST12-2/ BKL87 and the *hrdB* internal control probe (for a complete description see Fig. 3.19) are indicated as well as the detected *hrdB* transcript.


(A)
KGI10
JST12-2
GTG
GTG
ATG
bldG
sco3550
BKL87


(B)
40 µg
G A T C 24 48
Probe+S1
Probe-S1
Full Length JST12-2/BKL87 Probe (+ non-homologous tail)
Full Length hrdB Probe (+ non-homologous tail)
hrdB transcript

transcript was present (Figure 3.20 B). Because internal S1 nuclease protection analysis relies on distinction between full-length protection and probe-probe reannealing, the size of the non-homologous extension is an important consideration. Therefore the experiment was repeated using the KGI10/BKL87 probe. The *sco3550* complementary region of primer KGI10 was identical to that of JST12-2 and contained sequence homology to the *sco3550* predicted ATG start codon but did not contain a non-homologous extension (Fig. 3.20 A). However when this probe was used, *sco3550* transcripts were not detected despite the appearance of *hrdB*-specific signals. It was hypothesized that *sco3550* expression was at levels too low to allow transcript detection (Figure 3.20 B). This did not discount the possibility that *sco3550* transcript is highly unstable, processed in some manner or exhibiting a high degree of RNA secondary structure such that it completely impedes S1 nuclease transcript analysis.

It should be noted that 3' S1 nuclease mapping of *sco3550* transcripts was also attempted using probe KGI11/ KGI12 that included 87 bp upstream and 154 bp downstream of the predicted UAG stop codon (also containing a non-homologous extension on the end downstream of the stop codon). The probe was labeled on its 3' end by *Bam*HI digestion of the probe and filling in with radiolabelled dCTP, however, hybridizing transcripts were not detected. Although this experiment was hard to interpret because of the lack of an internal control, a failure in 3' S1 nuclease mapping may be indicative of the similar problems involved in 5' mapping or may be representative of the occurrence of 3' exonuclease digestion of the transcript.

3.6.2 Primer extension analysis of *sco3550* transcripts

Since the S1 nuclease experiments had failed to detect *sco3550* transcript, primer extension was examined as an alternative technique for assessment of *sco3550* transcript levels and transcriptional start site identification. However, a variety of end-labelled primers were used which were complementary to different *bldG-sco3550* intergenic sequence and this led to the corresponding identification of multiple *sco3550* transcriptional start sites. This inconsistency in *sco3550* transcript and transcription start site identification based on the primer used in these experiments was attributed to predicted secondary structure formation in the *bldG-sco3550* intergenic region (or the 5' *sco3550* UTR). Therefore, primer annealing in the predicted *sco3550* 5' UTR may alter the secondary structure, much like the binding of small RNAs can induce RNA secondary structure rearrangements, and possibly terminate the primer extension reaction at different end-points.

3.7 Biochemical characterization of recombinant MBP-SCO3550

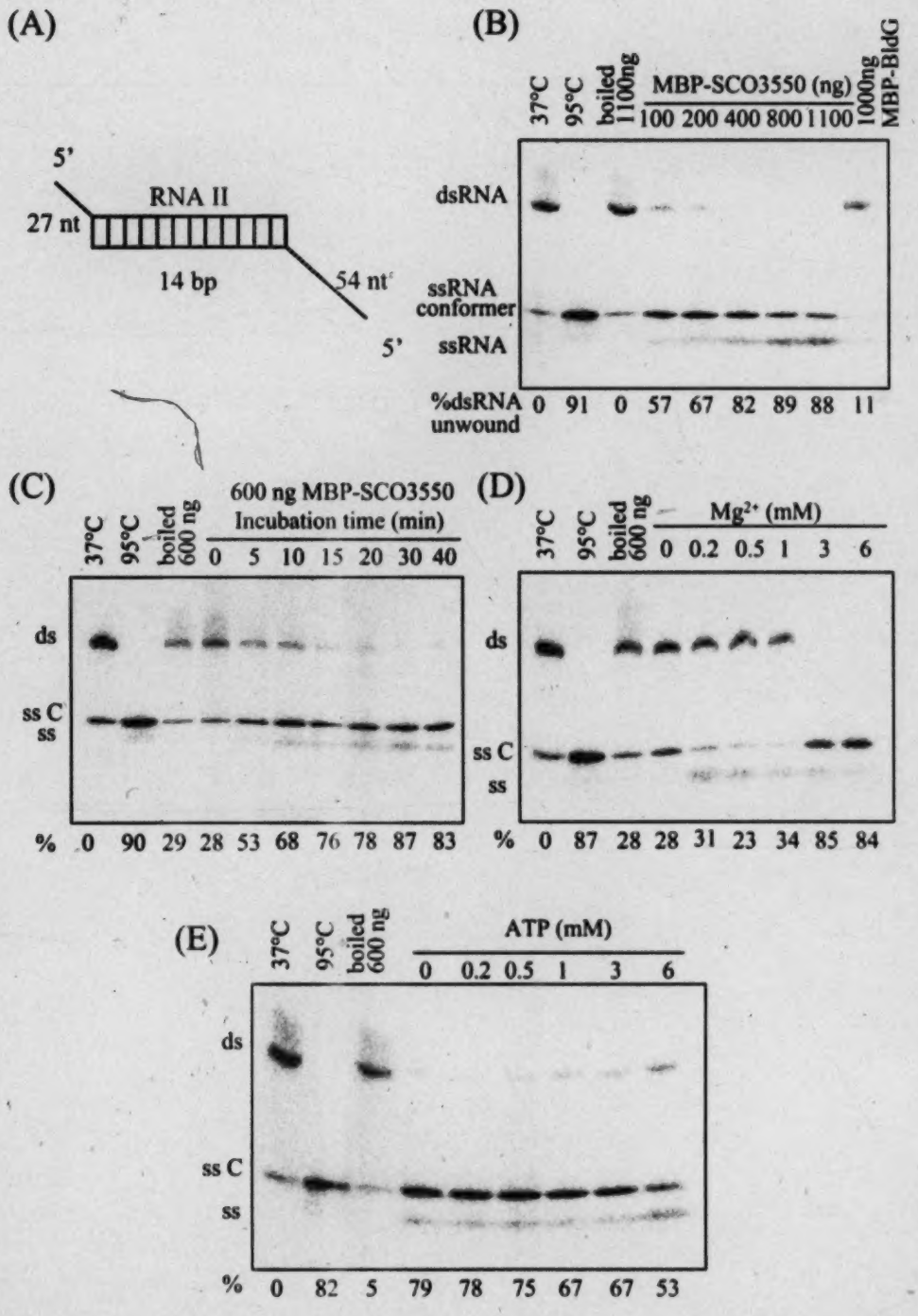
Predicted amino-acid sequence analysis of SCO3550 (Stoehr, 2001) revealed similarity to the superfamily 2 of helicases, which are a large group of RNA helicases that exhibit RNA-dependent ATP hydrolysis activity and ATP-dependent RNA-unwinding activity in order to unwind RNA secondary structure. The presence in SCO3550 of 7 conserved motifs shared with DexH-box RNA helicases, together with the similar spacing between those motifs, indicated a clustered catalytic core and experiments were performed to characterize the *in vitro* RNA helicase activity of recombinant SCO3550 protein.

3.7.1 *in vitro* RNA unwinding assays

Using the method described by Pause and Sonenberg (1995) and Yu and Owttrim (2000), *in vitro* RNA unwinding assays were undertaken with MBP-SCO3550 fusion protein (Pause and Sonenberg, 1992). Assays were performed using an artificial, partially double-stranded RNA substrate with either 5' or 3' single-stranded overhangs (generously provided by the Owttrim lab). In this type of assay, the double-stranded region of the substrate is created by intermolecular base-pairing between two single-stranded RNA molecules, one of which is radiolabeled. MBP-SCO3550 was added to the radiolabeled substrate under standard assay conditions (see Section 2.7.1). Reaction products were separated by electrophoresis and dsRNA and ssRNA species were examined by virtue of differing mobilities on the gel and were visualized and quantified by phosphorimager and ImageQuant analysis.

Initial substrate loading of RNA helicases appears to be dependant upon single-stranded RNA binding and progressive movement along the substrate in either a 5'→3' or 3'→5' polarity in order to promote unwinding of the dsRNA region (Hesson et al., 2000). Therefore, the RNAII artificial duplex substrate was utilized to assess the 5'→3' *in vitro* RNA unwinding activity of SCO3550. Structure of the artificial duplex RNA substrate, RNAII, formed by the intermolecular base-pairing of a non-radioactive 41 nt and radiolabeled 68 nt single-stranded RNA molecule, consists of 5' 27 and 54 nt single-stranded overhangs and a 14 bp double-stranded core (Yu and Owttrim, 2000) (Figure 3.21 A). Recombinant MBP-SCO3550 (see Section 3.2) was added to the RNAII substrate, in varying concentrations and from at least three independent fusion protein preparations, under standard assay conditions and was incubated at 37 °C for 30 minutes.

Figure 3.21 5'→3' unwinding activity of MBP-SCO3550. RNA unwinding assays were performed using standard assay conditions with 600 ng (amount which showed maximum substrate unwinding) of recombinant MBP-SCO3550 and incubation at 37 °C for 30 minutes unless otherwise stated. (A) The structure of the artificial, partially double-stranded RNA substrate, RNAII. RNA unwinding assays performed under various conditions are shown in Panels (B)-(E). Note that the dsRNA, ssRNA molecule and potential ssRNA conformers are indicated as well as the % dsRNA unwound. The % unwound was calculated using ImageQuant analysis where the negative control, the protein-free 37 °C sample, was used as the 0% dsRNA unwound standard. As a positive control for RNAII duplex destabilization the sample was boiled and incubated in the absence of recombinant MBP-SCO3550 and is indicated as 95 °C. MBP-BldG was also used as a negative control in panel (B) to rule out the pssibility that any MBP-tagged proteins could be stimulated to effect RNA duplex unwinding. All reactions contain a boiled protein negative control to ensure that the protein component of the added MBP-SCO3550 sample was responsible for unwinding. (B) The reactions were performed using various concentrations of recombinant MBP-SCO3550, (C) various incubation times at 37 °C, (D) various levels of Mg^{2+} and (E) various levels of ATP as indicated. The products were analyzed by 10 % SDS-PAGE and subsequent exposure to a phosphorscreen.


(A)
5'
RNA II
27 nt
14 bp
54 nt
(B)
37°C
95°C
boiled 1100ng
MBP-SCO3550 (ng)
100 200 400 800 1100
1000ng MBP-BldG
dsRNA
ssRNA conformer
5' ssRNA
%dsRNA unwound
0 91 0 57 67 82 89 88 11
(C)
600 ng MBP-SCO3550
Incubation time (min)
37°C
95°C
boiled 600 ng
0 5 10 15 20 30 40
ds
ss C ss
% 0 90 29 28 53 68 76 78 87 83
(D)
37°C
95°C
boiled 600 ng
Mg2+ (mM)
0 0.2 0.5 1 3 6
ds
ss C ss
% 0 87 28 28 31 23 34 85 84
(E)
37°C
95°C
boiled 600 ng
ATP (mM)
0 0.2 0.5 1 3 6
ds
ss C ss
% 0 82 5 79 78 75 67 67 53

The percentage of ds RNAII substrate unwound is indicated and based upon the ds substrate alone being representative of 0 % unwinding. After separation of the products by 10% SDS-PAGE and exposure to a phosphorimaging plate screen, it was clear that recombinant SCO3550 could mediate complete duplex RNAII unwinding when added to the reaction at concentrations higher than 400 ng as indicated by a nearly complete shift from double-stranded to single-stranded RNA species (Figure 3.21 B). Note that this is quite a high amount of protein in comparison to other RNA helicases and may suggest low *in vitro* unwinding activity. For example, complete duplex unwinding was observed at 20 ng levels of CrhC by Owtrimm and Yu (2000). Two single-stranded RNA species were visualized on the gel and the upper ssRNA, which was present in minor amounts in the protein free control, was believed to be a conformer with a slight ds nature due to foldback and base-pairing of the ssRNA. Due to its presence at background levels throughout the course of the experiment, it was possible that this conformer could form in the absence of the helix unwinding or destabilizing activity of SCO3550 and was resistant to unwinding by the addition of the recombinant protein. SCO3550 unwinding activity was observed relative to negative controls for RNA unwinding, such as a protein free control and a sample using boiled recombinant protein. Also, to confirm that MBP did not have any nascent RNA unwinding activity, MBP-BldG was used in the assay and showed low background levels of ds destabilization. It was also observed that increasing the length of incubation time prompted a corresponding increase in the amount of partial duplex RNAII converted to ssRNA, and a 30 minute incubation time was necessary to facilitate complete RNAII unwinding by the addition of 600 ng of MBP-SCO3550 (Figure 3.21 C). It was concluded that recombinant MBP-SCO3550 possesses 5' RNAII

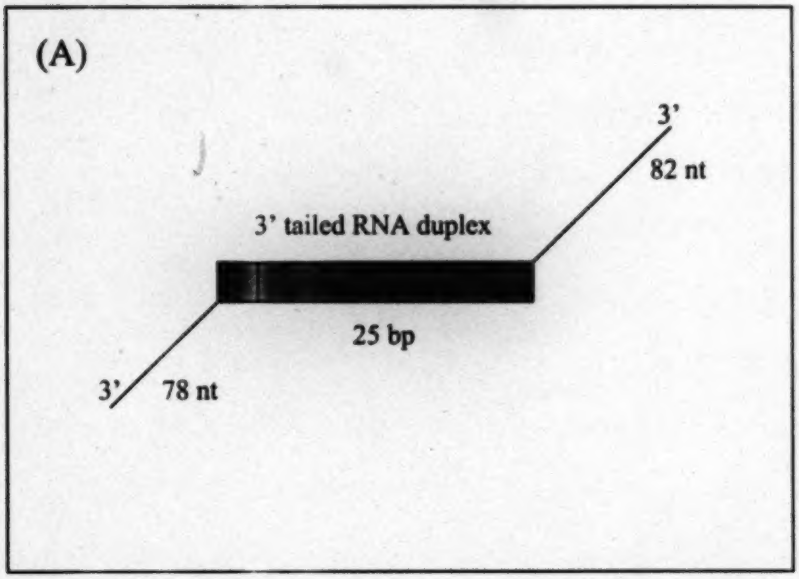
unwinding activity albeit to a lesser degree than many other putative RNA helicases examined using this protocol (Kujat-Choy, 2000; Yu and Owttrim, 2000; Chamot et al., 2004). It should also be noted that to verify that recombinant MBP-SCO3550 was responsible for the duplex unwinding observed in the assay, co-immunoprecipitation experiments with anti-SCO3550 polyclonal antibodies were attempted to remove the protein from the reaction mixture. However, the treatment of protein samples with proteinA-sepharose after antibody addition failed to completely impede the unwinding reaction which was most likely owing to the incomplete removal of MBP-SCO3550 from the reaction. This could be ascribed to the lack of specificity for the antibodies prior to affinity purification and the experiments were not repeated using the affinity-purified antibody.

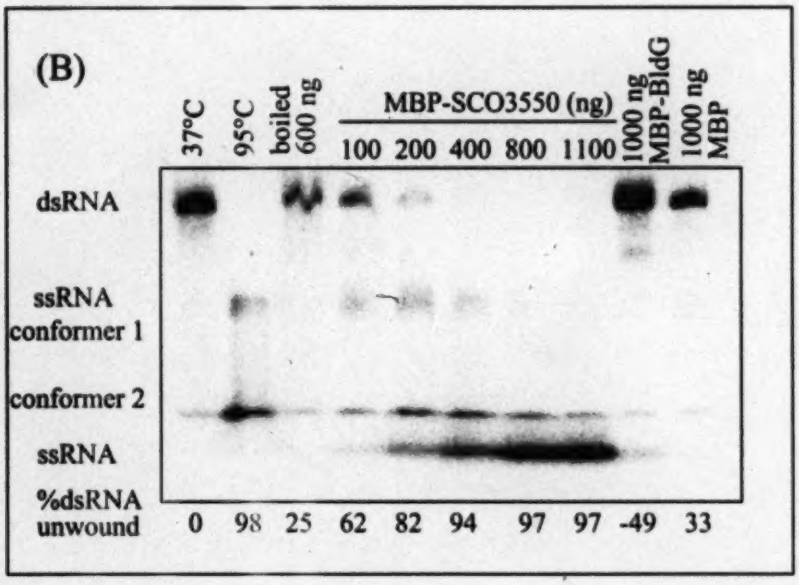
X-ray crystallographic studies that suggest ATP binding occurs through the first aspartate residue of the DECH (or prototypical DEAD) motif. ATP binding most likely occurs through aspartate interaction with gamma and beta phosphates mediated by a water molecule and Mg^{2+} (Hodgman, 1988; Kujat-Choy, 2000). Therefore the Mg^{2+}-dependence of recombinant MBP-SCO3550 RNAII unwinding was examined. The recombinant MBP-SCO3550 (600 ng) was incubated with RNAII under standard assay conditions except for the addition of varying amounts of $MgCl_2$. Due to the inability of MBP-SCO3550 to mediate conversion of RNAII from a ds to ss form when Mg^{2+} levels were less than 3 mM it was concluded that SCO3550 *in vitro* duplex unwinding activity has an absolute requirement for Mg^{2+} (Figure 3.21 D). It was also hypothesized that Mg^{2+} dependence is necessary for ATP binding and subsequent hydrolysis to enable enzyme processivity.

Finally, the ATP dependence of recombinant SCO3550 RNAII unwinding was examined by incubation of 600 ng recombinant protein with RNAII under standard assay conditions in the presence of varying levels of ATP. Unexpectedly, SCO3550 unwinding activity was not ATP dependent as the recombinant protein mediated nearly complete unwinding of the RNAII partial duplex in the absence of ATP addition (Figure 3.21 E). The high amount of SCO3550 required to enable complete dsRNA destabilization, combined with its ATP-independence, predicts that SCO3550 may be a specific-RNA binding protein or RNA chaperone. These enzymes can act to passively destabilize RNA secondary structure and prime the RNA for a protein interaction or conformational change (Jones et al., 1996). Consequently, the Mg^{2+} dependence of SCO3550-mediated RNAII destabilization would be less likely to be involved in ATP binding and rather could stabilize or aid SCO3550 binding and destabilization of RNA secondary structure. Alternatively, the MBP-SCO3550 fusion protein may be tightly bound to ATP that may be co-purified from the *E. coli* host strain. This would create the illusion of ATP-independence in a reaction lacking exogenous ATP. A large amount of protein added, which is already bound to ATP, may provide enough enzyme to promote complete destabilization of the substrate.

A separate artificial, partially double-stranded RNA substrate was also utilized that consisted of 3' 78 and 82 nt single-stranded overhangs and a 25 bp double-stranded core (again, generously provided by the Owtrimm lab) (Yu and Owttrim, 2000) (Figure 3.22 A). Originally the 3' tailed RNA duplex was used in RNA unwinding assays to determine the predicted polarity of MBP-SCO3550 RNA loading and the progression of enzyme unwinding, but the absence of bonafide ATP-dependent RNA helicase activity

Figure 3.22 3'→5' unwinding activity of MBP-SCO3550. RNA unwinding assays were performed using standard assay conditions with 600 ng of recombinant MBP-SCO3550 and incubated at 37 °C for 30 minutes unless otherwise stated. (A) The structure of the artificial 3' single-stranded partial duplex substrate. (B) Various concentrations of recombinant SCO3550 were incubated with the RNA duplex and products were analyzed by 10 % SDS-PAGE and subsequent exposure to a phosphorscreen. Note that the dsRNA, ssRNA molecule and various potential ssRNA conformers are indicated as well as the % dsRNA unwound. The % unwound was calculated using ImageQuant analysis and the negative control, the protein-free 37 °C sample, was used as the 0% dsRNA unwound standard. As a positive control for duplex destabilization the sample was boiled and incubated in the absence of recombinant MBP-SCO3550 and is indicated as 95 °C, while MBP-BldG and MBP alone were used as negative controls to determine if MBP affected RNA duplex unwinding. Boiled MBP-SCO3550 was also used as a negative control to test that the protein component of the MBP-SCO3550 sample added to the reaction was responsible for substrate unwinding.


(A)
3'
82 nt
3' tailed RNA duplex
25 bp
3'
78 nt


(B)
37°C
95°C
boiled
600 ng
MBP-SCO3550 (ng)
100 200 400 800 1100
1000 ng MBP-BldG
1000 ng MBP
dsRNA
ssRNA conformer 1
conformer 2
ssRNA
%dsRNA unwound
0 98 25 62 82 94 97 97 -49 33

for the MBP-SCO3550 negates the relevance of this question. Rather, the 3' tailed RNA duplex substrate was used to test the potential for recombinant MBP-SCO3550 to bind and destabilize the duplex. In this respect, varying amounts of recombinant SCO3550 protein were added to the 3' tailed RNA duplex under standard assay conditions. Subsequent examination by 10 % SDS-PAGE and exposure to a phosphorimaging plate screen revealed that complete duplex destabilization could occur at greater than 400 ng of added recombinant protein, similar to the observed amounts required in the 5' duplex unwinding assay (Figure 3.22 B). Also, MBP did not appear to have an effect on 3' dsRNA destabilization as MBP-BldG and MBP alone displayed very little decrease in double-stranded RNA. Extra species were visualized and assumed to be additional single-stranded conformers due to the extended length of the substrate and a longer region of base-pairing than in the 5' RNAII artificial substrate.

As a follow-up experiment, it was hypothesized that *in vitro* RNA duplex destabilization by recombinant SCO3550 must necessitate a degree of duplex binding and therefore, reaction products were observed on native 10% polyacrylamide gels where protein binding would be observed by a shift or smear in the radiolabeled substrate. However, even in reactions where formaldehyde was used to theoretically cross-link the protein to nucleic acid, a shift in the radiolabeled duplex was not observed (not shown). This was a baffling finding and currently it is not fully understood why protein binding of the RNA substrate is not readily observed unless the binding itself is extremely transient or unstable within an *in vitro* setting.

3.8 HIS_6-SCO3550 Overexpression and Binding Studies in *Streptomyces*

3.8.1 Protein Overexpression and Purification in *Streptomyces*

It is an interesting faculty of RNA helicases that sees them often in association with co-factors or coupled to macromolecular machines to exact their cellular function (von Hippel and Delagoutte, 2003). Therefore, overexpression studies were performed in *S. coelicolor* wild-type strain M600 in order to identify potential protein binding partners of the putative helicase protein (and for future experiments involving identification of SCO3550 *in vivo* RNA targets). Recombinant SCO3550 was overexpressed from pAU336 in *S. coelicolor* as a 6x histidine-tagged protein under the control of the thiostrepton inducible p*tipA* promoter (Table 2.4).

Design of the overexpression plasmid involved the use of the pIJ6902 derivative, pAU331, containing the thiostrepton inducible p*tipA* promoter, and integrates into the *Streptomyces* chromosome at the *attB* site. Initial cloning entailed *NdeI* restriction enzyme digestion of a 60 bp histidine (6x) tag-encoded DNA fragment, which also contained a thrombin cleavage site for liberation of the tag. The gel-purified 60bp fragment was provided by Linda Bui of the Leskiw lab. pAU331, containing *sco3550* cloned under the control of the p*tipA* promoter in pIJ6902, was then digested with *Nde*I and subsequently dephosphorylated with 10 U of shrimp alkaline phosphatase to hinder vector re-circularization in the ligation reaction. Ligation of the *Nde*I-digested pAU331 and the 60 bp HIS_6 fragment, proceeded in a 10 µL final volume and the ratio of vector to insert was slightly biased towards the vector in order to impede the formation of insert concatemers. Potential transformants were screened by *Nde*I restriction digestion to liberate the 60 bp fragment, and also by PCR amplification, using primers LB1, which

annealed at the 5' portion of the HIS_6 fragment, and BKL87, which annealed 225 bp into *sco3550*. Out of the 6 transformants chosen, all 6 revealed a liberated fragment of approximately 60 bp in size upon *NdeI* digestion and 5 out of the 6 showed PCR amplification of the appropriate product size. Transformant 6 showed a series of higher molecular weight PCR products in addition to the desired product and it was assumed that this was representative of possible cloning of HIS_6 concatemers into pAU331 (data not shown). Finally, 3 of the positive transformants were chosen for automated sequencing performed with primer BKL87, and a single transformant containing the correctly cloned HIS_6 fragment upstream of the translational start of SCO3550 was designated pAU336. pAU336 was transferred into *E. coli* strain ET12567/ pUZ8002 by transformation and then transferred by conjugation into *S. coelicolor* M600.

Initial HIS_6-SCO3550 induction experiments were performed on plate culture and phenotypes were examined relative to the amount of thiostrepton inducer that was added to the medium (Stoehr, 2001). The phenotypic effect of induction of recombinant SCO3550 was examined on R2YE plates containing apramycin (for the selection of the integrated plasmid) and a range of thiostrepton concentrations (0-100 μg/mL) to induce the expression of HIS_6-SCO3550. Growth was only very slightly delayed by HIS_6-SCO3550 induction in comparison to the control strain containing insert-free pIJ6902, and the same degree of growth defect was seen regardless of the amount of inducer added to the test strain. To verify that HIS_6-SCO3550 was actually being induced, recombinant SCO3550 was further examined by Western analysis using affinity-purified anti-SCO3550 antibodies. Figure 3.23 confirms induction of HIS_6-SCO3550 in cell-free extracts purified from strains of *S. coelicolor* containing pAU336 and grown on R2YE

Figure 3.23 Western analysis of a plate culture induction of HIS_6-SCO3550. Crude extracts from strain M600 + pAU336 were isolated after growth for 48 hr on R2YE solid medium overlayed with cellophane disks and containing either 0, 10, 30 or 100 μg/mL of thiostrepton inducer. Fifty micrograms of each sample were separated by 6 % SDS-PAGE. After transfer to a PVDF membrane, samples were probed with affinity purified anti-SCO3550 polyclonal antibodies followed by a secondary hybridization with Horseradish peroxidase-conjugated secondary antibody at a dilution of 1:10 000. Detection was performed by a 1 minute incubation with Enhanced chemiluminescence reagents and exposure to Kodak film for 1-10 minutes. Kaleidoscope pre-stained standards were used to estimate the size of the detected bands and also to monitor the efficacy of electroblot transfer of proteins to the PVDF membrane. The recombinant HIS_6-SCO3550 band is indicated (i) and native SCO3550 (ii) are indicated.

(A)

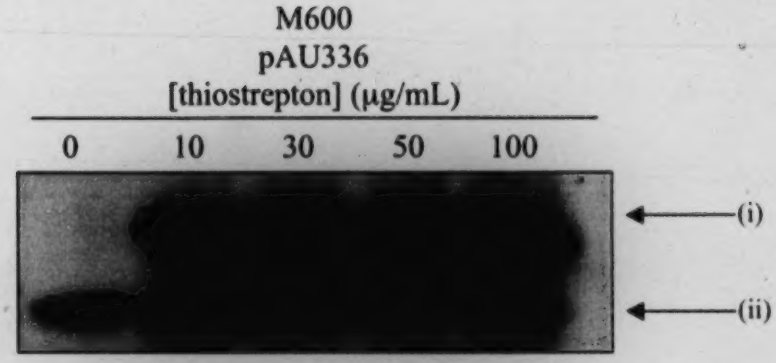


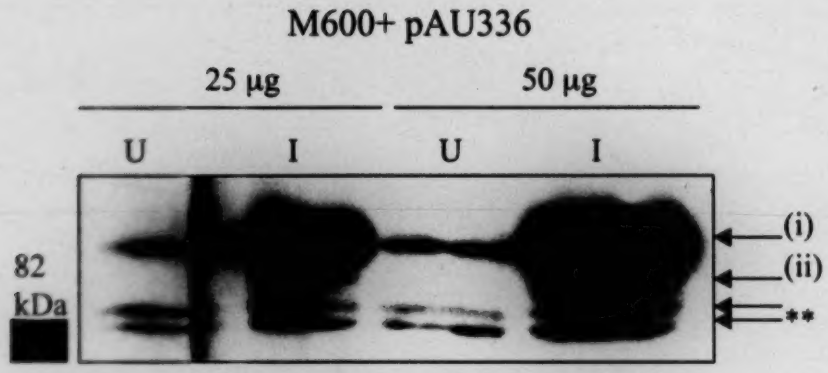
plate medium with the addition of varying concentrations of thiostrepton ranging from 0-100 μg/mL. Furthermore, the HIS_6-SCO3550 was visualized by western analysis using 6% SDS-PAGE and, as expected of a protein with a highly charged tag (Bignell et al., 2000), found to migrate at a higher position that wild-type SCO3550. Recombinant SCO3550 was observed to be absent in the absence of thiostrepton induction and fully induced when thiostrepton concentrations exceeded 10 μg/mL. Because the induction was carried out for more than 48 hours, only the smaller putative isoform of native SCO3550 appears on the immunoblot (see Figure 3.16).

For a larger scale induction of HIS_6-SCO3550, *S.coelicolor* M600 + pAU336 was grown in 50 mL R2YE broth and induced at a high OD with the addition of a final concentration of 30 μg/mL thiostrepton. Cell extracts from both an uninduced and induced culture were harvested after a 3 hour incubation at 30 °C and 25 and 50 μg of cell-free extract was separated by 6 % SDS-PAGE. Recombinant SCO3550 expression was then examined by western analysis (Figure 3.24 A). Evidently, HIS_6-SCO3550 was present in minor amounts in the uninduced crude extract, which suggests a degree of leaky expression from the p*tipA* promoter in the absence of thiostrepton induction in liquid culture only (refer to Figure 3.23). Expression in the induced crude extract was very high in comparison to the level of the lower molecular weight native isoforms of wild-type SCO3550 (see Section 3.5 for explanation of the two bands).

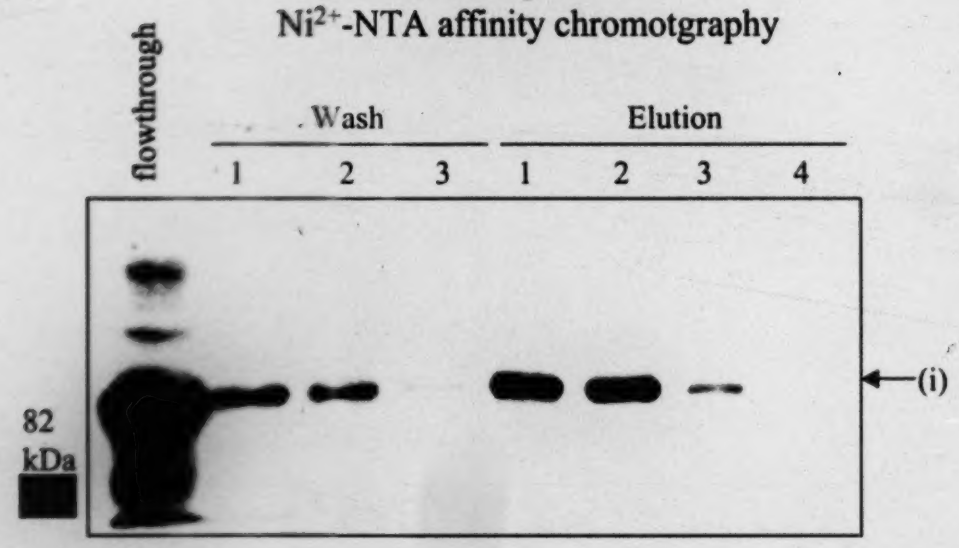
Initial purification of HIS_6-SCO3550 was attempted using Qiagen Ni^{2+}-NTA micro-spin columns for rapid purification. However, this technique was developed to facilitate rapid purification of highly overexpressed recombinant proteins and unsuccessful attempts to isolate HIS_6-SCO3550 were attributed to the relatively low level

Figure 3.24 Liquid culture induction of HIS_6-SCO3550 expression and Ni^{2+}-NTA affinity chromotography. *Streptomyces* M600 + pAU336 was grown in 50 mL R2YE broth to a high O.D., induced with 30 µg/mL final concentration of thiostrepton and incubated at 30 °C for 3 hours in order to overexpress HIS_6-SCO3550. (A) Crude extracts from both induced samples and the uninduced control were isolated and 25 and 50 µg of each sample, as indicated, were separated by 6 % SDS-PAGE and then subjected to western analysis using affinity purified anti-SCO3550 antibodies (1:10 000 dilution). (B) Purification of recombinant SCO3550 was performed with a Ni^{2+}-NTA Batch Purification Protocal where 10 mg of induced crude extract was incubated with 500 µL of Ni^{2+}-NTA resin at 4 °C for 1 hour while gently shaking. The resin-extract mixture was loaded onto a 5 mL column and allowed to flowthrough. Subsequently, the resin was washed three times with 4 mL 1Xwash buffer and eluted in four separate fractions with 500 µL of 1Xelution buffer (containing 250 mM imidazole). Samples were TCA precipitated, redissolved in 25-30 µL 2x SDS- loading dye and 15% of the total volume was examined by 6 % SDS-PAGE and western analysis. For the western analysis shown in both (A) and (B) samples were probed with affinity purified anti-SCO3550 polyclonal antibodies. A secondary hybridization was performed with Horseradish peroxidase-conjugated secondary antibody at a dilution of 1:10 000. Detection was performed by a 1 minute incubation with Enhanced chemiluminescence reagents and exposure to Kodak film for 2 minutes. Kaleidoscope pre-stained standards were used as a size marker and position of the 82 kDa marker is shown. The recombinant His_6-SCO3550 band (i) as well as a potential recombinant protein degradation band (ii) and the predicted native SCO3550 species, are also indicated (**).

(A)


M600+ pAU336
25 µg
50 µg
U
I
U
I
82
kDa
(i)
(ii)
**

(B)


M600+ pAU336
Ni2+-NTA affinity chromotgraphy
flowthrough
Wash
Elution
1
2
3
1
2
3
4
82
kDa
(i)

induction in comparison with *E. coli* overexpression systems, and to the short binding time associated with the columns. Subsequent purification of of HIS_6-SCO3550 was tested using a Ni^{2+}-NTA resin Batch Purification Protocol by Novagen. This procedure allowed the purification of recombinant protein from a high volume of crude extract and longer resin binding times were believed to contribute to a higher degree of HIS_6-SCO3550 purification. Briefly, 10 mg of crude extract isolated from the thiostrepton induced M600 + pAU336 strain was incubated with Ni^{2+}-NTA resin at 4 °C for at least 1 hour while gently mixing. The resin was then washed and HIS_6-SCO3550 was eluted, and the flowthrough, wash and elution fractions were saved for analysis. Figure 3.24 B shows the western analysis of 15% of the isolated Ni^{2+}-NTA affinity chromotography fractions. The major portion of recombinant SCO3550 was present in the flowthrough fraction, which may be rectified by longer protein-resin binding times and an increase in resin volume. However, recombinant SCO3550 was also present in the elution fractions. This was in contrast to affinity chromotography performed on the uninduced crude extract samples that showed no detectable hybridization by western analysis (not shown). The remaining 85 % of the isolated samples were analyzed by 6% SDS-PAGE followed by Coomassie blue staining and revealed a very faint band that corresponded to the migration of HIS_6-SCO3550 as detected by western analysis (data not shown). This was also accompanied by several faint background bands that were also present in the uninduced sample. Therefore, it was determined that affinity chromotography required some refinements to increase the amount of recombinant protein isolated in order to eventually be able to identify potential binding partners by Coomassie Brilliant Blue staining and mass spectroscopy. For instance, considerations kept in mind for later

chemical crosslinking studies (see Section 3.8.2-3.8.3) were the necessity for longer resin binding times and several repetitions of washing to remove protein bound to the column non-specifically.

3.8.2 *in vivo* Chemical crosslinking experiments

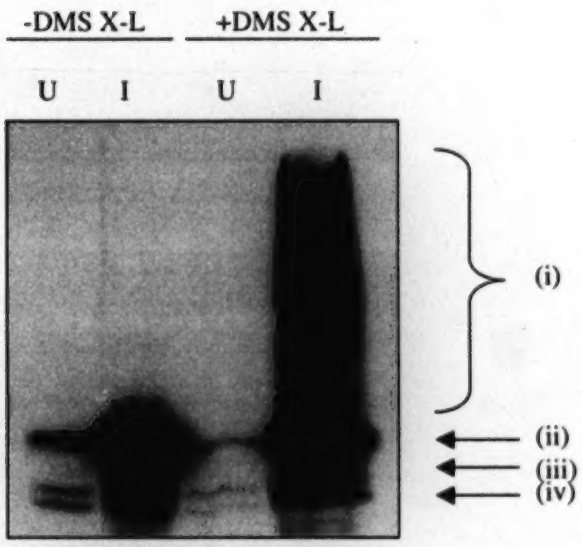
Preliminary chemical crosslinking experiments were performed to covalently attach amino groups of HIS_6-SCO3550 to amino groups of potential protein binding partners. Crosslinking was expected to enrich the amount of complexes, or protein-protein interactions that could be isolated from the cell and purified with Ni-NTA affinity chromotography. Initially, *in vivo* chemical crosslinking was performed to stabilize any putative HIS_6-SCO3550 protein complexes prior to cell lysis which may normally endanger the integrity of potential complexes or abolish the normal cellular context of the protein in question. A homobifunctional lysine-specific crosslinking reagent DMS (3,3'-dimethyl suberimidate) (MJS BioLynx) was used in the *in vivo* crosslinking reaction in combination with DMSO, in order to increase the permeability of the cell to the cross-linking reagent (Watty et al., 1997). Formaldehyde was also used in this experiment to enhance protein-protein crosslinking, but more importantly to promote crosslinking between HIS_6-SCO3550 and its potential nucleic acid substrates. The use of formaldehyde was based on the inverse assumption that the identification of nucleic acid binding substrates is often complemented with protein-protein crosslinking in order to maintain the correct cellular context of the protein (Kurdistani and Grunstein, 2003). Therefore, an interaction with a nucleic acid substrate may, by the same virtue, be

important for the cellular context of HIS_6-SCO3550 and its interaction with other proteins' or involvment in large molecular assemblies.

Strains were prepared for *in vivo* DMS crosslinking similar to the liquid induction of HIS_6-SCO3550 descibed above. *S.coelicolor* M600 + pAU336 was grown in 4X 50 mL R2YE broth and induced at a high OD with the addition of thiostrepton to a final concentration of 30 μg/mL. The mycelia were pelleted and washed with 10 mL of ice-cold PBS buffer and resuspended in 10 mL PBS. DMS was added to 10 mmol final concentration, DMSO to 0.25% and formaldehyde to 1% and the reaction mixture was incubated for 1 hour at room temperature with very slow mixing (to avoid frothing but ensure some gentle mixing). An uninduced control culture was also subjected to the same crosslinking conditions and uncrosslinked controls were also prepared for both, HIS_6-SCO3550 uninduced and induced cultures. The mycelia in the reaction mixtures were pelleted and washed in PBS buffer and resuspended in a small volume of PBS containing protease inhibitor and the suspensions were then sonicated on ice for at least 2X 15 seconds to lyse the cells. The resulting ruptured cell suspension was centrifuged at 14 000 rpm for 10 minutes at 4 °C, the supernatant containing the soluble cell extract was decanted and the protein contents of the extracts were determined.

Prior to attempted Ni-NTA affinity chromotography, 50 μg of cell extract from both the uncrosslinked controls and DMS crosslinked samples were examined by western analysis using affinity-purified anti-SCO3550 antibodies (Figure 3.25). HIS_6-SCO3550 appears to be overexpressed in both induced uncrosslinked and cross-linked samples as expected, while there is a background level of HIS_6-SCO3550 in the uninduced samples as previously observed. Interestingly, the induced DMS crosslinked sample contains a

Figure 3.25 *in vivo* protein-protein crosslinking with DMS. *Streptomyces* M600 containing pAU336 was grown in 50 mL R2YE broth to a high O.D., and HIS_6-SCO3550 expression was induced with thiostrepton (30 μg/mL final concentration) and incubated at 30 °C for 3 hours. The *in vivo* crosslinking reaction was performed by the addition DMS to 10 mmol, DMSO to 0.25% and formaldehyde to 1% to mycelia resuspended in PBS buffer. The reaction mixture was incubated for 1 hour at room temperature with slow mixing. The mycelia were lysed by sonication and crude extracts from both uninduced (U) and induced (I) uncrosslinked (-DMS X-L) and uninduced (U) and induced (I) crosslinked (+DMS X-L) samples were isolated and 50 μg of each sample was separated by 6 % SDS-PAGE and then subjected to western analysis. Upon transfer to a PVDF membrane, samples were probed with affinity purified anti-SCO3550 polyclonal antibodies (1:10 000 dilution) followed by a secondary hybridization with Horseradish peroxidase-conjugated secondary antibody at a dilution of 1:10 000. Detection was performed by a 1 minute incubation with Enhanced chemiluminescence reagents and exposure to Kodak film for 2 minutes. (i) indicates the high molecular weight smear possibly representative of His_6-SCO3550 containing complexes that are crosslinked to varying degrees, while, the recombinant His_6-SCO3550 band (ii) as well as the putative larger (iii) and smaller (iv) SCO3550 isoforms, are also indicated.


-DMS X-L
+DMS X-L
U
I
U
I
(i)
(ii)
(iii)
(iv)

smear of higher molecular weight bands that are absent from the other samples. In addition the presence of the high molecular weight bands appear to diminish the amount of monomeric HIS_6-SCO3550 present. This suggests that this smear may represent complexes containing His_6-SCO3550 or multimerization of HIS_6-SCO3550 itself. It was suggested that because DMSO is not completely efficient at promoting cell permeability to DMS (personal communication, Brenda Leskiw) that a high molecular weight complex containing His_6-SCO3550 may have varying degrees of intermolecular bond formation between separate complexes. A complex may then undergo varying degrees of intermolecular disruption which would lead to a large number of altered migration positions when examined upon a denaturing gel and could yield this observed smear. To test this hypothesis, *in vitro* chemical crosslinking was attempted to allow saturation of the cell extract with a chemical crosslinking reagent.

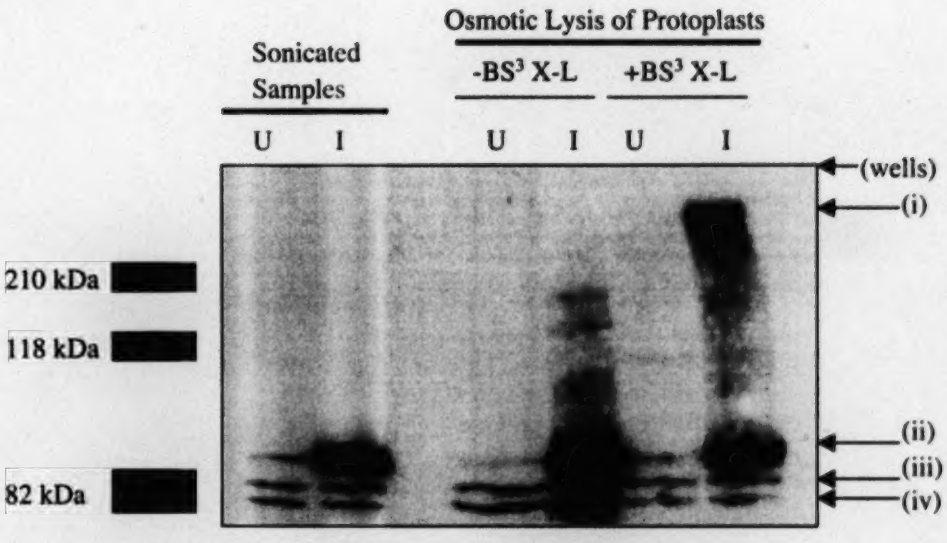
3.83 *in vitro* Chemical crosslinking experiments

The efficiency of *in vitro* chemical crosslinking relies upon the gentle lysis of the cell and maintainance of protein-protein interactions until they can be covalently attached by adding the crosslinking reagent to the soluble cell extract. BS^3 [bis(sulfosuccinimidyl) suberate], a lysine-specific crosslinking reagent, (Kurdistani and Grunstein, 2003), was used due to its efficacy in *in vitro* crosslinking experiments performed by other lab members (personal communication Brenda Leskiw and Archana Paranshar). Strains were prepared for *in vitro* BS^3 crosslinking similar to the liquid induction used for *in vivo* HIS_6-SCO3550 DMS-crosslinking as described above. *S.coelicolor* M600 + pAU336 was grown in 4X 50 mL R2YE broth and induced at a high OD with the addition of

thiostrepton at a final concentration of 30 μg/mL. Note that a sample was left uninduced and aliquots of both uninduced and induced samples were not chemically crosslinked to provide necessary controls for comparison to the induced crosslinked sample. As a gentle lysis procedure, prior to BS^3 crosslinking, cells were protoplasted and osmotically lysed by resuspension in a small volume of crosslinking reaction buffer. Cell extract samples were then crosslinked for a 1 hr incubation at 30 °C by addition of BS^3 crosslinking reagent, in combination with DNase, to degrade chromosomal DNA and reduce the viscosity of the solution, and protease inhibitor. The crosslinking reaction was stopped by the addition of lysine to 1% of the total reaction volume and the reaction mixture was centrifuged at 13 000 rpm for 10 minutes at 4 °C and the protein content of the soluble extract was determined.

To assess the presence of potential HIS_6-SCO3550 protein-protein interactions, 40 μg of uninduced and induced uncrosslinked and crosslinked extracts were examined by western analysis using affinity-purified anti-SCO3550 antibodies (Figure 3.26). A distinct high molecular weight band was observed in the induced crosslinked sample that was not observed in the control samples and its presence appeared to diminish the amount of the monomeric HIS_6-SCO3550 present in the sample. In addition, the absence of this band in the induced uncrosslinked sample indicates that chemical crosslinking is necessary to observe this species. By the same virtue, the absence of the band in the uninduced crosslinked sample suggests that by overexpressing HIS_6-SCO3550, we may be fostering the occurrence of a SCO3550 interaction with a potential binding partner. Consequently, *in vitro* BS^3 crosslinking appeared to allow identification of a potential high molecular weight complex that involves the participation of HIS_6-SCO3550. The

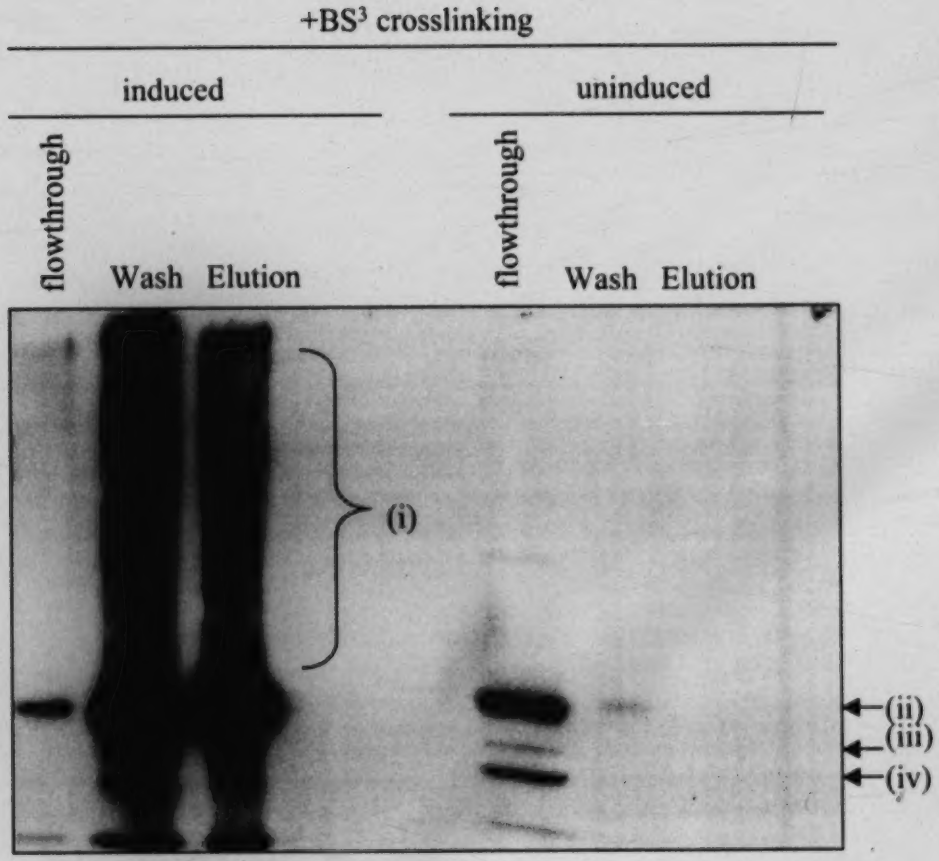
Figure 3.26 *in vitro* protein-protein crosslinking with BS^3. *Streptomyces* M600 containing pAU336 was grown in 50 mL R2YE broth to a high O.D., and HIS_6-SCO3550 expression was induced with thiostrepton (30 μg/mL final concentration) and incubation at 30 °C for 3 hours. The cell mycelia were gently lysed by protoplasting followed by osmotic lysis after resuspension in a 2-4 mL volume (depending upon biomass) of crosslinking reaction buffer. Cell extract samples were then crosslinked for a 1 hr incubation at 30 °C by adding the BS^3 crosslinking reagent, in combination with DNase and protease inhibitor. The crosslinking reaction was stopped by the addition of lysine to 1% of the total reaction volume and the reaction mixture was centrifuged at 13 000 rpm for 10 minutes and the protein content of the soluble extract was determined. Crude extracts (40 μg) from both uninduced (U) and induced (I) uncrosslinked (-BS^3 X-L) and uninduced (U) and induced (I) crosslinked (+BS^3 X-L) samples were isolated and separated by 6 % SDS-PAGE and then subjected to western analysis using affinity purified anti-SCO3550 antibodies. Note that extracts prepared by sonication of mycelia which expression of HIS_6-SCO3550 had either been uninduced (U) or induced (I) were also included as a negative and positive control for protein expression (labeled as sonicated samples). Western analysis was carried out using the standard procedure (See Figure 3.25) after separation of the proteins by 6 % SDS-PAGE. (i) indicates a potential high molecular complex containing HIS_6-SCO3550 (and possibly native SCO3550), while, the recombinant HIS_6-SCO3550 band (ii) as well as the putative larger (iii) and smaller (iv) native SCO3550 isoforms are also indicated.


Osmotic Lysis of Protoplasts
Sonicated Samples
-BS³ X-L
+BS³ X-L
U
I
U
I
U
I
(wells)
(i)
210 kDa
118 kDa
82 kDa
(ii)
(iii)
(iv)

complex appears larger than the largest Kaleidoscope standard which is 210 kDa. It is of course, a possibility that this high molecular weight band is simply a multimeric form of HIS_6-SCO3550, however, superfamily 2 RNA helicases are believed to only interact as dimers (see introduction).

In order to determine the nature of the high molecular weight band, purification by Ni-NTA affinity chromotography was necessary. A Ni-NTA Batch Purification Protocol was performed similar to that described in section 3.8.1, however, a larger volume of Ni-NTA resin (2mL) was used. Also, it was suggested that periodic, gentle mixing when binding to the resin may enhance the binding effeciency of HIS_6-SCO3550 as the HIS tag- Ni-NTA interaction may be weakened by a low representation of HIS_6-SCO3550 in the high molecular weight complex and/or partial steric occlusion of tag binding sites by the complex (personal communication, Troy Locke). Both, uninduced and induced BS^3 crosslinked samples were subjected to Ni-NTA affinity chromotography and 15 % of the flowthrough, wash and eluate fractions was examined by western analysis (Figure 3.27) and the remaining 85 % was examined by Coomassie Brilliant Blue staining. Figure 3.27 clearly shows that a high degree of HIS_6-SCO3550 monomer and the potential complex are being eluted in the wash fraction and yet, both species also appear to be highly represented in the pooled eluate sample. This is in contrast to uninduced crosslinked samples which display some leaky expression of monomeric HIS_6-SCO3550 in the flowthrough fraction. Also, the potential complex appears as a smear which may possibly be due to the high level of crude extract used in this experiment (4 mg) and protein overloading on the gel. Unfortunately, the Coomassie Blue stained gel showed no discernible protein bands in either the uninduced or induced eluate lanes and

Figure 3.27 Ni-NTA affinity chromotography of BS^3 crosslinked samples. Purification of monomeric HIS_6-SCO3550 and the potential high molecular weight complex was attempted using a Ni^{2+}-NTA Batch Purification Protocal detailed in section 3.8.1. Crude extract (4 mg) from both HIS_6-SCO3550 induced and uninduced strains, which had been subjected to BS^3 crosslinking, were incubated with 2 mL of Ni^{2+}-NTA resin at 4 °C for 1 hour with gentle, periodic mixing. The resin-extract mixture was loaded onto a 5 mL column and allowed to flowthrough. Subsequently, the resin was washed three times with 4 mL 1Xwash buffer and eluted in four separate fractions with 500 µL of 1Xelution buffer (containing 250 mM imidazole). Wash and Eluate samples were pooled and TCA precipitated, redissolved in 25-30 µL 2x SDS- loading dye and 15% of the total volume was examined by 6 % SDS-PAGE and western analysis. Upon transfer to a PVDF membrane, samples were probed with affinity purified anti-SCO3550 polyclonal antibodies followed by a secondary hybridization with Horseradish peroxidase-conjugated secondary antibody at a dilution of 1:10 000. Detection was performed by a 1 minute incubation with Enhanced chemiluminescence reagents and exposure to Kodak film for 2 minutes. (i) indicates the high molecular weight smear possibly representative of His_6-SCO3550 containing complexes, while, the recombinant His_6-SCO3550 band (ii) as well as the putative larger (iii) and smaller (iv) native SCO3550 isoforms, are also indicated.


+BS3 crosslinking
induced
uninduced
flowthrough
Wash
Elution
flowthrough
Wash
Elution
(i)
(ii)
(iii)
(iv)

there was also no difference between the banding pattern observed in the induced versus the uninduced wash sample lanes (data not shown). The sensitivity of the western analysis in contrast to the Coomassie blue staining likely accounts for the failure to detect bands on the stained gel. Therefore, at this stage sufficient amounts of monomeric HIS_6-SCO3550, or of the potential complex involving HIS_6-SCO3550, have not been purified by Ni-NTA affinity chromotography for mass spectroscopy. It has been suggested that because the number of cellular proteins in the size range of the potential high molecular weight complex are very few, that for future experiments, the amount of the complex may be enriched by size exclusion chromotography and sufficient levels of the pure complex may be visualized by 2-D gel electrophoresis thereby circumventing the use of affinity chromotography altogether.

Chapter 4:

Discussion

4. Discussion

Over the course of this study, characterization of *sco3550* and its encoded putative RNA helicase was attempted in order to clarify the nature of the gene product and define its role in the developmental program of *S. coelicolor*. Interest in *sco3550* was originally piqued because of its genomic location being situated closely adjacent (with predicted overlapping promoters) to *bldG* and divergently expressed relative to the *bldG* locus. Therefore, the implicit goal of this project was to increase the level of data and experience pertaining to the study of *sco3550* and RNA helicases in general, so as to be poised for the next generation of experiments aimed at defining the cellular role for *sco3550* by analyzing specific RNA targets and protein binding partners.

Early work on *sco3550* by J. Stoehr revealed the gene to be an intractable subject to many forms of experimentation (Stoehr, 2001). Attempts to assess the gene expression profile of *sco3550* by numerous techniques and construction of a (single) disruption mutant were never fully realized and it was hypothesized that *sco3550* has a very low level of expression and may in fact be an essential gene. With this in mind, I set out to further characterize *sco3550* and perhaps circumvent some of the difficulties that had been experienced in the past by trying either separate methodologies or variations of those that had already been attempted.

Initially, the protein expression profile of SCO3550 was examined by western analysis using anti-SCO3550 polyclonal antibodies that had been generated using MBP-SCO3550 fusion protein as antigen. In the early analyses, SCO3550 was shown to be expressed at a nearly constant level throughout the life-cyle of *S. coelicolor* with a peak

at 24-36 hours however, the lack of a true positive or negative *in vivo* control for the western analysis together with the high degree of background cross-hybridization, hindered any definite conclusions from being drawn. To reduce the amount of background when using anti-SCO3550 antibodies, affinity purification based on the specificity of the polyclonal antibodies to the MBP-SCO3550 fusion protein, was performed. When western analysis of SCO3550 was repeated using the affinity purified antibodies, the level of background hybridization was significantly lowered and the specific SCO3550 hybridizing band was revealed to be a potential doublet. Using a lower percentage gel, the single band that was originally considered to represent the native form of SCO3550 was resolved into two separate bands with very unique expression patterns. Through the construction of a positive control where *sco3550* was overexpressed in a putative *bldG*3b/*sco3550* double mutant strain, the lower band of the doublet was confirmed to be a specific SCO3550 hybridizing band and the upper band was then assumed to be a protein cross-reacting with the anti-SCO3550 antibodies. The lower band was found to be present during early vegetative growth, reached peak levels at 24 hr, the time corresponding to aerial hyphal formation, and then decreased to a very low level by 48 hr when sporulation was complete.

Even more interestingly, the expression pattern of both bands in the doublet was disrupted in the *bldG*1DB in-frame deletion mutant, lending support to the proposed link between *bldG* and *sco3550*, and also strongly suggesting that both bands of the doublet were forms of SCO3550. This dependence of SCO3550 expression upon either an intact *bldG* wild-type gene or functional BldG suggests that an epistatic arrangement may exist between *bldG* and *sco3550*, where *bldG* is responsible for regulation of *sco3550*.

However, without the benefit of examining *bldG* transcript or protein levels in a *sco3550* disruption mutant or overexpression strain, it has not escaped our notice that *sco3550* dependence upon the *bldG* locus may be part of a feedback regulatory system where upregulated or activated SCO3550 may then impose negative or positive regulation on *bldG* or *orf3*.

In addition to the western analyses described above, several lines of evidence support the conclusion that the observed doublet is in fact two separate isoforms of SCO3550. To begin, BLASTP searches of the *S. coelicolor* gene database (http://www.sanger.ac.uk/Projects/S_coelicolor/) were performed to identify potential homologues or proteins similar enough to exhibit a strong immuno cross-reactivity with SCO3550. When the entire SCO3550 sequence was searched against the *S. coelicolor* genomic database, only two putative ATP-dependent RNA helicases were uncovered that had molecular weights similar to SCO3550. SCO664 has 33% identity and SCO4096 has 26 % identity to SCO3550, and have deduced molecular weights of 80.2 and 81.4 kDa respectively. Although relatively close, these molecular weights are too small to confuse with the 86 kDa SCO3550 protein, especially given the resolving power of the 6% PAG. Because of the large C terminal domain present in SCO3550, which presumptively gives the enzyme specificity and cellular context through protein-protein interactions (Rocak and Linder, 2004), it is possible that proteins containing similar domains are present in *S. coelicolor* and could result in cross-reaction with the anti-SCO3550 antibodies. These proteins may be masked in the BLASTP searches performed with the entire amino acid sequence of SCO3550 due to the high number of conserved RNA helicase motifs found in the N-terminal domain of the protein. Yet, a search of the C-terminal domain against

the genomic database revealed no strong homologs and the proteins showing weak similarity to SCO3550 were nowhere in the size range of 86 kDa. Finally, because antibodies were generated and affinity-purified against the entire MBP-SCO3550 fusion protein, a BLASTP search of MBP against the genomic database was performed and only revealed putative maltose and sugar binding proteins that all lie within the size range of 40-50 kDa. Thus, with the exception of cleavage of much larger RNA helicases or multiple post-translational modifications of smaller putative RNA helicases that are similar to SCO3550, there are no predicted proteins within the size range of SCO3550 in *S. coelicolor* that provide reasonable targets for cross-reactivity with MBP-SCO3350 affinity-purified antibodies.

Clearly, the presence of two separate forms of SCO3550 in *S. coelicolor* may provide the key to understanding many of the impasses experienced in conducting experiments and interpretation of results concerning *sco3550*. The possibility of at least two enzymatic activites or a variety of target specificities for the separate isoforms of SCO3550 might fit with the biphasic gene expression patterned observed by Stoehr (2001). If true, the peak of transcript levels observed at 15 hr would result in expression of the large isoform of SCO3550 and the mRNA peak at 24-36 hr would result in expression of the smaller isoform. In this scenario the two isoforms would be expected to exhibit different roles in the organism. The paradigm shift from the genome project-predicted GTG translational start site to the designated ATG start codon (Stoehr, 2001) may have skewed the design of our experiments to favor the smaller of the predicted SCO3550 isoforms. For example, the results from the *bldG*3b/*sco3550* double mutant strain harboring pAU333, which is responsible for overexpressing SCO3550 protein in

the presence of thiostrepton inducer, showed overexpression of only the lower band of the doublet. In retrospect this is exactly what would have been expected because only the region of *sco3550* including and downstream of the designated ATG translational start codon was included in the construct. Coding sequence for the N-terminal sequence of the putative larger isoform of the protein was of course left absent because it was not known to be relevant. This is also a possible explanation for the failure of the conditional mutagenesis experiment as only the coding sequence downstream of the genome predicted ATG start codon was included in the pAU332 plasmid containing the inducible extra copy of *sco3550* and may have excluded vital upstream sequence that would be necessary to complement the disruption of the wild-type gene. It is unclear at this time if there are dual promoters for the gene, or separate translational start sites from the same transcript, but the possiblity of proteolytic processing does appear low due to the presence of two in-frame suitable translational start codons. Further experimentation must be undertaken to address these questions. One possibility is to use immunoprecipitation experiments and mass-spectroscopy (MS) to identify the purified 87 and 86 kDa species. Preliminary immunoprecipitation experiments initially showed little success but are currently underway with the affinity-purified antibody. Also, overexpression of the SCO3550 protein from the upstream predicted GTG codon must be performed to show an induction of both hybridized bands in the doublet; however, this may be difficult due to the lack of a true null *sco3550* background to distinguish weakly induced species from the native species also present. Finally, the blockage in examining gene expression of *sco3550* must be surpassed in order to confirm the transcriptional start

site, test the possibility of multiple promoters, transcript proccessing and to examine the potential for *sco3550* to be encoded as a leaderless transcript.

Mutagenesis of *sco3550* was also attempted using a PCR-targetted mutagenesis protocol. Although a fraction of the chromosomal complement of *S. coelicolor* was disrupted, the strain appeared to maintain a small fraction of wild-type chromosomes that could not be cleared from the strain. Similar results have been shown for *div*IVA in *S. coelicolor* (Flardh, 2003). The author found that strains that were predicted to be double cross-over exconjugants and disruption mutants for *div*IVA based upon the nature of antibiotic resistance, still maintained the wild-type allele and it was presumed that the kanamycin resistance gene had been lost from the integrated cosmid. There was no mention of the possibility that the strain may contain a mixture of wild-type and mutant chromosomes as only chromosomal PCR of the individual wild-type or mutant allele was used for verification. The *div*IVA gene was subsequently reasoned to be an essential gene, and a conditional mutant was constructed using a similar strategy to the one employed in this study.

In retrospect, it is clear that there are many technical problems with the experimental design of the attempted conditional mutagenesis that may have contributed to the failure of obtaining a conditional mutant isolate. First of all, the designed experiment depends upon *sco3550* being an essential gene in *S. coelicolor* because if the conditional mutant has no visible phenotype there is no way to identify the uninduced strain as a *sco3550* chromosomal null mutant. As already discussed, Southern analysis is not sensitive enough to detect wild-type chromosomes when present as a small fraction of a multigenomic mixture. Western analysis would also have its disadvantages as there is

no way to differentiate SCO3550 protein as expressed from a wild-type copy of the gene that has resisted disruption or as exogenous protein expressed from the inducible extra copy. One possibility to differentiate wild-type from exogenous protein would be to tag the inducible copy of the protein. For example, a HIS_6-tag may not affect the *in vivo* activity of the protein but could be distinguished from wild-type protein based on differing mobilities in SDS-PAGE.

The technical difficulties described above are compounded with evidence that in some systems the p*tipA* promoter has low level leaky expression in the absence of thiostrepton induction and in this case, the depletion experiments to isolate and determine the phenotype of a conditional mutant would be ineffective (personal communication B. Leskiw, see Section 3.2.2). This was an issue in the design of the *div*IVA conditional mutant, however, the strain showed a mutant phenotype in the absence of inducer even though a background level of *div*IVA expression occurred as a result of leaky *ptipA* expression (Flardh, 2003). In this case the level of *divIVA* expression was far less than the level necessary for wild-type growth and division. For genes with predicted low transcript abundance, leaky expression from the uninduced p*tipA* promoter would preclude complete depletion of the gene product and therefore of assessing the null mutant phenotype. Construction of a conditional mutant may also require repeated rounds of platings and generation of spore stocks in the presence of thiostrepton induction to ensure that all wild-type copies of the gene are disrupted or segregated out of the strain.

An additional problem experienced in the attempt to construct a conditional *sco3550* mutant was that western analysis of the thiostrepton induced strains did not

show high induction relative to wild-type protein expression levels (data not shown). Therefore, expression from pAU332 may not produce SCO3550 protein at sufficient levels (or timing) to create the atmosphere under which a conditional disruption could occur. Although a conditional mutant was not obtained, the observed difficulties in *sco3550* mutagenesis and instability of the *bldG*3b/*sco3550* putative double mutant predict that *sco3550* may be playing an essential role in *S. coelicolor* and future refinements to a conditional mutagenesis protocol may yet substantiate this hypothesis.

In addition to trying to establish a role for *sco3550* in morphological differentiation and antibiotic production in *S. coelicolor*, this study also set out to show that SCO3550 functions as a RNA helicase. The biochemical activity of the MBP-SCO3550 fusion protein was examined using *in vitro* enzyme assays and was determined to lack ATP-dependent RNA unwinding, while still displaying the ability to destabilize the partial duplex substrate when high amounts of protein were added to the reaction. There appears to be a clear discrepancy between the presence of conserved motifs purported to be involved with RNA helicase activity and the actual *in vitro* activity of recombinant MBP-SCO3550. However, many putative RNA helicases, as predicted by conserved motifs, are recalcitrant to *in vitro* artificial duplex unwinding assays. Examples include the *E. coli* putative helicase CsdA (Jones et al., 1996) and mammalian p68 (Hirling et al., 1989), which both display ATP-independent dsRNA destabilizing activity. The inability of these enzymes to show RNA unwinding alone or unwinding coupled to ATPase activity, is suggested to be due to the absence of specific RNA substrates, co-factors and/or post-translational modifications that may stimulate the activity of the putative RNA helicase (Chen et al., 2002). Therefore, the biological

function of SCO3550 cannot be so easily assigned from a single *in vitro* dsRNA unwinding assay and RNA helicase activity should not be readily discounted. It is has already been shown by western analysis that there are two possible isoforms of SCO3550 possibly generated by separate domains of transcription and translation or through various post-translational modifications and both the larger and smaller isoforms may have distinct biological roles and separate *in vitro* activities. In this study only the smaller molecular weight isoform was examined for potential RNA destabilizing activity and it will be necessary to examine the biochemical nature of the larger molecular weight isoform.

It has been shown that a large group of RNA helicases require a co-factor or binding partner to enhance RNA helicase activity by stabilizing substrate or ATP interactions or by other uncharacterized mechanisms. For example, the prototypical RNA helicase eIF4-A, eukaryotic initiation factor 4-A, which is involved in translation initiation, requires ancillary factor eIF4-B to significantly stimulate eIF4-A RNA helicase activity (Rozen et al., 1990; Rogers et al., 1999). Another example of the need for protein-protein interactions to stimulate helicase activity is the yeast protein Dbp5, which absolutely requires the co-immunoprecipitation of an uncharacterized binding partner in order to exhibit *in vitro* ATP-dependent RNA unwinding activity (Tseng et al., 1998). Consequently, immunoprecipitation of SCO3550 or identification of binding partners through other means such as, chemical cross-linking of recombinant SCO3550 overexpressed in *Streptomyces*, may allow a more biologically relevant reconstitution of SCO3550 enzyme activity within *in vitro* assays. Of course, the possibility cannot also be discounted that the MBP N-terminal tag may be affecting the structure and activity of

the protein especially since the proposed catalytic core resides in the N-terminus of the protein might be sterically blocked by the MBP. Therefore future enzyme assays should be conducted with FPLC purified native protein or with fusion protein containing a smaller tag, such as HIS_6-tagged recombinant SCO3550.

Another possibility for the failure to show ATP-dependent RNA unwinding may have been due to the co-purification of ATP with the MBP-SCO3550 fusion protein from the *E. coli* overexpression strain. In this scenario, even when ATP was not added to the reaction mixture, there may have been enough ATP-bound protein present to promote duplex destabilization. This artifact may be exaggerated when a large excess of protein must be added due to low efficiency of the reaction, as was the case for MBP-SCO3550, and under these conditions hydrolyzed ATP may be playing a role in dissociation of the protein from the substrate rather than directing enzymatic conformational changes. It would be possible to examine possible co-purification of MBP-SCO3550 with ATP by adding an excess of the non-hydrolyzable analogue ATP-γS to the reaction mixture. An ATP exchange experiment would show a reduction in duplex destabilization corresponding to an increase in the addition of ATP-γS if MBP-SCO3550 was indeed using the energy from ATP hydrolysis to drive duplex destabilization.

Based upon the behaviour of a putative RNA helicase in an *in vitro* unwinding assay setting, three models have been postulated to describe potential enzyme binding, and the mode, efficiency and processivity of the enzymatic mechanism (Rocak and Linder, 2004). The 'on and off' model pertains to helicases with low activity *in vitro* and which are required in molar excesses to catalyze duplex destabilization. In this model, the association and dissociation of an RNA helicase with single-stranded RNA may

destabilize downstream duplex RNA. Double-stranded RNA undergoes rapid and constant thermal breathing at duplex ends and the authors reason that an association of an RNA helicase close to a region of thermal denaturation may further destabilize the duplex and block reannealing. In the 'on and off' model, ATP hydrolysis allows active dissociation of the protein from the nucleic acid and repeated rounds of ATP binding, nucleic acid binding, and ATP hydrolysis allow the protein component to be recycled. More efficient RNA helicases may unwind double-stranded RNA by the 'Translocation model'. Using ATP hydrolysis, the RNA helicase may translocate along single-stranded RNA and when encountering regions of duplex thermal breathing would impede duplex re-annealing. Therefore, this is considered to be passive unwinding as the initial force of unwinding is due to nascent thermal denaturation but would still be more efficient than the 'on-off' model as the reaction would not rely upon a fortuitous association at the site of thermal denaturation. The third model, 'Translocation and unwinding', applies to highly efficient and processive DNA and RNA helicases. The model predicts that the helicase would use energy from ATP hydrolysis to translocate and actively destabilize the RNA duplex, presumably by forcing the RNA backbone to assume an unstable configuration. This model would probably not describe most RNA helicases, which are believed to only unwind short RNA duplexes, but may relate to more processive RNA helicases, such as, ribonucleoproteinases (RNPases) that require processivity to destabilize RNA-protein interactions. Clearly, an *in vitro* mechanism of duplex destabilization relating to MBP-SCO3550 must be similar to the 'on-off' model where association with the RNA substrate is leading to duplex denaturation, however, it is still

questionable whether ATP is playing a role in the reaction and the *in vivo* mechanism of action is still clouded.

In the last part of this thesis work, experiments were performed that utilized expression of a histidine (6x)-tagged version of SCO3550 from a chromosomal integrant in wild-type *S. coelicolor* and chemical crosslinking studies were done to examine the occurrence of protein-protein interactions involving the recombinant protein. It was observed that HIS_6-SCO3550 was possibly involved in a high molecular weight complex that could only be visualized in cell-extracts that had been initially induced for recombinant protein expression and subjected to chemical crosslinking. However, preliminary experiments to purify the complex were unsuccessful and how to effectively purify this complex in order to identify its nature by mass spectroscopy remains unresolved but may be solved by scaling-up the procedure.

The occurrence of a putative RNA helicase in a large complex is a common theme in the study of RNA helicases (Rocak and Linder, 2004) and supports the variety of models predicting the potential cellular function of SCO3550 postulated by J. Stoehr [see Discussion, (Stoehr, 2001)]. Briefly, the three most probable macromolecular complexes that SCO3550 may be associating with are the degradosome, the ribosome, or the RNA polymerase holoenzyme assembly. Interaction with any of the mentioned assemblies could hypothetically exert an influence upon (at least) the *bldG* locus. As already mentioned there are predicted to be three major regions of RNA secondary structure formation in the overlapping *sco3550* and *bldG* locus: within the intergenic region of *sco3550-bldG*, the intergenic region of *bldG-orf3* and in the downstream untranslated region of *orf3*. RNA secondary structure re-arrangements and unwinding by SCO3550,

perhaps associated within a macro-molecular complex, may facilitate and regulate the activity of the complex at the *bldG* locus or other targeted loci, be it at the level of transcript turn-over, expression or translation. We still have little data substantiating the claim that SCO3550 effects regulation of the *bldG* locus and it is a possibility that the reverse could be true. Figure 3.18 A suggests that SCO3550 protein expression is dependent upon an intact *bldG* gene or functional protein and that SCO3550 may play a role in developmental that is downstream of the *bldG* locus. Of course, interaction of SCO3550 with a complex may be instituting finer controls over the activity of the degradosome, transcription or translation apparatus or even ribosomal biogenesis that may affect a large subset of genes. It would be very interesting to correlate the nature of the cellular function of SCO3550 with the scope of its effects and examine if the interaction of SCO3550 with a macro-molecular complex can shift its function from a vegetative to a developmental focus.

Several developments in this study should facilitate future work that seeks to examine potential targets and binding-partners of SCO3550 with the ultimate goal to define its cellular function throughout the *S. coelicolor* life cycle of growth and differentiation. For instance, affinity purified anti-SCO3550 antibodies should greatly increase the specificity of the entire antibody complement and enhance the ability of immunoprecipiation techniques to isolate native forms of SCO3550. Also, the creation of a *S. coelicolor* strain that overexpresses a histidine tagged version of SCO3550 will be pivotal for identification and purification of protein-binding partners as well as identification of RNA substrates. Consequently, this study has led to the door-step of two probable methods for co-purifying RNA substrates that interact with SCO3550: co-

immunoprecipitation and affinity chromotography of recombinant SCO3550. Of course, an unbiased approach to identify these RNA substrates will now be the central difficulty but is not an insurmountable problem.

Although the probability of creating a viable disruption strain of *sco3550* appears unlikely, the creation of a strain in this study that overexpresses *sco3550* from a chromosomal integrating plasmid may serve a purpose in identifying the function of SCO3550. Furthermore, an isogenic strain that overexpresses SCO3550 will provide the next best thing to a mutant as a control for experiments that assess changes in the proteome by 2-D gel electrophoresis and may identify targets that are influenced by the overexpression of SCO3550. It is still something of an unknown factor if SCO3550 exerts its control over a wide range of genes, or embodies a very narrow focus on targets such as the *bldG* locus, and how this contributes to the many physiological and morphological changes inherent in the developing mycelium.

Chapter 5:

References

Alting-Mees, M. A. and Short, J. M. (1989). pBluescript II: gene mapping vectors. *Nucleic Acids Res* **17**, 9494.

Ainsa, J.A., Parry, H.D. and Chater, K.F.: A response regulator-like protein that functions at an intermediate stage of sporulation in *Streptomyces coelicolor* A3(2). Mol Microbiol 34 (1999) 607-19.

Ainsa, J.A., Ryding, N.J., Hartley, N., Findlay, K.C., Bruton, C.J. and Chater, K.F.: WhiA, a protein of unknown function conserved among Gram-positive bacteria, is essential for sporulation in *Streptomyces coelicolor* A3(2). J Bacteriol 182 (2000) 5470-8.

Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J. and Hopwood, D.A.: Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417 (2002) 141-7.

Bibb, M.J. and Buttner, M.J.: The *Streptomyces coelicolor* Developmental Transcription Factor σ^{BldN} Is Synthesized as a Proprotein. J Bacteriol 185 (2003) 2338-45.

Bibb, M.J., Molle, V. and Buttner, M.J.: σ^{BldN}, an extracytoplasmic function RNA polymerase sigma factor required for aerial mycelium formation in *Streptomyces coelicolor* A3(2). J Bacteriol 182 (2000) 4606-16.

Bignell, D.R., Lau, L.H., Colvin, K.R. and Leskiw, B.K.: The putative anti-anti-sigma factor BldG is post-translationally modified by phosphorylation in *Streptomyces coelicolor*. Submitted (2003).

Bignell, D.R., Warawa, J.L., Strap, J.L., Chater, K.F. and Leskiw, B.K.: Study of the *bldG* locus suggests that an anti-anti-sigma factor and an anti-sigma factor may be involved in *Streptomyces coelicolor* antibiotic production and sporulation. Microbiology 146 (2000) 2161-73.

Boddeker, N., Stade, K. and Franceschi, F.: Characterization of DbpA, an *Escherichia coli* DEAD box protein with ATP independent RNA unwinding activity. Nucleic Acids Res 25 (1997) 537-45.

Bradford, M.M.: A rapid and sensitive method for the quantification of microgram quantities of proteins utilizing the principle of protein dye-binding. Anals of Biochemistry 72 (1976) 248-254.

Bruton, C.J., Plaskitt, K.A. and Chater, K.F.: Tissue-specific glycogen branching isoenzymes in a multicellular prokaryote, *Streptomyces coelicolor* A3(2). Mol Microbiol 18 (1995) 89-99.

Bystrykh, L.V., Fernandez-Moreno, M.A., Herrema, J.K., Malpartida, F., Hopwood, D.A. and Dijkhuizen, L.: Production of actinorhodin-related "blue pigments" by *Streptomyces coelicolor* A3(2). J Bacteriol 178 (1996) 2238-44.

Caruthers, J.M., Johnson, E.R. and McKay, D.B.: Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. Proc Natl Acad Sci U S A 97 (2000) 3080-85.

Caruthers, J.M. and McKay, D.B.: Helicase structure and mechanism. Curr Opin Struct Biol 12 (2002) 123-33.

Castrillon, D.H., Quade, B.J., Wang, T.Y., Quigley, C. and Crum, C.P.: The human VASA gene is specifically expressed in the germ cell lineage. Proc Natl Acad Sci U S A 97 (2000) 9585-90.

Chaconas, G. and van de Sande, J.H.: 5'-32P labeling of RNA and DNA restriction fragments. Methods in Enzymology 65 (1980) 75-85.

Chakraburtty, R. and Bibb, M.: The ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. J Bacteriol 179 (1997) 5854-61.

Chamot, D., Colvin, K.R., Kujat-Choy, S.L. and Owttrim, G.W.: RNA structural rearrangement via unwinding and annealing by the cyanobacterial RNA helicase, CrhR. J Biol Chem [Epub ahead of print] (2004).

Champness, W.: Actinomycete development, antibiotic production, and phylogeny: questions and challenges. In: Shimkets, L.J. (Ed.), Prokaryotic Development. ASM Press, Washingtion, DC, 2000a, pp. 11-31.

Champness, W.: Actinomycete develpment, antibiotic production, and phylogeny: questions and challenges. ASM Press, 2000b.

Champness, W.C.: New loci required for *Streptomyces coelicolor* morphological and physiological differentiation. J Bacteriol 170 (1988) 1168-74.

Chater, K.F.: A morphological and genetic mapping study of white colony mutants of *Streptomyces coelicolor*. J Gen Microbiol 72 (1972) 9-28.

Chater, K.F.: Construction and phenotypes of double sporulation deficient mutants in Streptomyces coelicolor A3(2). J Gen Microbiol 87 (1975) 312-25.

Chater, K.F.: Taking a genetic scalpel to the *Streptomyces* colony. Microbiology 144 (1998) 1465-78.

Chater, K.F.: Regulation of sporulation in *Streptomyces coelicolor* A3(2): a checkpoint multiplex? Curr Opin Microbiol 4 (2001) 667-73.

Chater, K.F., Bruton, C.J., Plaskitt, K.A., Buttner, M.J., Mendez, C. and Helmann, J.D.: The developmental fate of *S. coelicolor* hyphae depends upon a gene product homologous with the motility sigma factor of *B. subtilis*. Cell 59 (1989) 133-43.

Chater, K.F. and Horinouchi, S.: Signalling early developmental events in two highly diverged Streptomyces species. Mol Microbiol 48 (2003) 9-15.

Chen, H., Lin, W., Tsay, Y., Lee, S. and Chang, C.J.: An RNA helicase, DDX1, interacting with poly(A) RNA and heterogeneous nuclear ribonucleoprotein K. J Biol Chem 277 (2002) 40403-9.

Cho, H.-S., Ha, N.-C., Kang, L.-W., Chung, K.M., Back, S.H., Jang, S.K. and Oh, B.H.: Crystal structure of RNA helicase from genotype 1b hepatitis C virus: a feasible mechanism of unwinding duplex RNA. J Biol Chem 273 (1998) 15045-15052.

Cho, Y.H., Lee, E.J., Ahn, B.E. and Roe, J.H.: SigB, an RNA polymerase sigma factor required for osmoprotection and proper differentiation of *Streptomyces coelicolor*. Mol Microbiol 42 (2001) 205-14.

Claessen, D., Wosten, H.A., van Keulen, G., Faber, O.G., Alves, A.M., Meijer, W.G. and Dijkhuizen, L.: Two novel homologous proteins of *Streptomyces coelicolor* and *Streptomyces lividans* are involved in the formation of the rodlet layer and mediate attachment to a hydrophobic surface. Mol Microbiol 44 (2002) 1483-92.

Coburn, G.A., Miao, X., Briant, D.J. and Mackie, G.A.: Reconstitution of a minimal RNA degradosome demonstrates functional coordination between a 3' exonuclease and a DEAD-box RNA helicase. Genes Dev 13 (1999) 2594-603.

Colley, A., Beggs, J.D., Tollervey, D. and Lafontaine, D.L.: Dhr1p, a putative DEAH-box RNA helicase, is associated with the box C+D snoRNP U3. Mol Cell Biol 20 (2000) 7238-46.

Davis, N.K. and Chater, K.F.: Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. Mol Microbiol 4 (1990) 1679-91.

Davis, N.K. and Chater, K.F.: The *Streptomyces coelicolor whiB* gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. Mol Gen Genet 232 (1992) 351-8.

Eccleston, M., Ali, R.A., Seyler, R., Westpheling, J. and Nodwell, J.: Structural and genetic analysis of the BldB protein of *Streptomyces coelicolor*. J Bacteriol 184 (2002) 4270-6.

Elliot, M., Damji, F., Passantino, R., Chater, K. and Leskiw, B.: The *bldD* gene of *Streptomyces coelicolor* A3(2): a regulatory gene involved in morphogenesis and antibiotic production. J Bacteriol 180 (1998) 1549-55.

Elliot, M.A., Bibb, M.J., Buttner, M.J. and Leskiw, B.K.: BldD is a direct regulator of key developmental genes in *Streptomyces coelicolor* A3(2). Mol Microbiol 40 (2001) 257-69.

Elliot, M.A., Karoonuthaisiri, N., Huang, J., Bibb, M.J., Cohen, S.N., Kao, C.M. and Buttner, M.J.: The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in *Streptomyces coelicolor*. Submitted (2003a).

Elliot, M.A. and Leskiw, B.K.: The BldD protein from *Streptomyces coelicolor* is a DNA-binding protein. J Bacteriol 181 (1999) 6832-5.

Elliot, M.A., Locke, T.R., Galibois, C.M. and Leskiw, B.K.: BldD from *Streptomyces coelicolor* is a non-essential global regulator that binds DNA as a dimer. Submitted (2003b).

Feinberg, A.P. and Vogelstein, B.: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anals of Biochemistry 132 (1983) 6-13.

Flardh, K.: Essential role of DivIVA in polar growth and morphogenesis in *Streptomyces coelicolor* A3(2). Molecular Microbiology 49 (2003) 1523-1536.

Flardh, K., Findlay, K.C. and Chater, K.F.: Association of early sporulation genes with suggested developmental decision points in *Streptomyces coelicolor* A3(2). Microbiology 145 (1999) 2229-43.

Fuller-Pace, F.V., Nicol, S.M., Reid, A.D. and Lane, D.P.: DbpA: a DEAD box protein specifically activated by 23S rRNA. Embo J 12 (1993) 3619-26.

Gorbalenya, A.E. and Koonin, E.V.: Helicases: amino acid sequence comparisons and structure-function relationships. Curr Opin Struct Biol 3 (1993) 419-29.

Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P. and Blinov, V.M.: Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res 17 (1989) 4713-30.

Gross, C.H. and Shuman, S.: The QRxGRxGRxxxG motif of the vaccinia virus DExH box RNA helicase NPH-II is required for ATP hydrolysis and RNA unwinding but not for RNA binding. J Virol 70 (1996) 1706-13.

Gust, B., Challis, G.L., Fowler, K., Kieser, T. and Chater, K.F.: PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for

biosynthesis of the sesquiterpene soil odor geosmin. Proc Natl Acad Sci U S A (2003).

Guthrie, E.P., Flaxman, C.S., White, J., Hodgson, D.A., Bibb, M.J. and Chater, K.F.: A response-regulator-like activator of antibiotic synthesis from *Streptomyces coelicolor* A3(2) with an amino-terminal domain that lacks a phosphorylation pocket. Microbiology 144 (1998) 727-38.

Hall, M.C. and Matson, S.W.: Helicase motifs: the engine that powers DNA unwinding. Mol Microbiol 34 (1999) 867-877.

Harasym, M., Zhang, L.H., Chater, K. and Piret, J.: The *Streptomyces coelicolor* A3(2) *bldB* region contains at least two genes involved in morphological development. J Gen Microbiol 136 (1990) 1543-50.

Havas, K., Flaus, A., Phelan, M., Kingston, R., Wade, P.A., Lilley, D.M. and Owen-Hughes, T.: Generation of superhelical torsion by ATP-dependent chromatin remodeling activities. Cell 103 (2000) 1133-42.

Herschlag, D.: RNA Chaperones and the RNA Folding Problem. J Biol Chem 270 (1995) 20871-74.

Hesketh, A., Chandra, G., Shaw, A., Rowland, J., Kell, D., Bibb, M. and Chater, K.: Primary and secondary metabolism, and post-translational modifications, as portrayed by proteomic analysis of *Streptomyces coelicolor*. Molecular Microbiology 46 (2002a) 917-932.

Hesketh, A.R., Chandra, G., Shaw, A.D., Rowland, J.J., Kell, D.B., Bibb, M.J. and Chater, K.F.: Primary and secondary metabolism, and post-translational protein modifications, as portrayed by proteomic analysis of *Streptomyces coelicolor*. Mol Microbiol 46 (2002b) 917-32.

Hesson, T., Mannarino, A. and M., C.: Probing the relationship between RNA-stimulated ATPase and helicase activities of HCV NS3 using 2'-O-methyl RNA substrates. Biochemistry 39 (2000) 2619-2625.

Hirling, H., Scheffner, M., Restle, T. and Stahl, H.: RNA helicase activity associated with the human p68 protein. Nature 339 (1989) 562-564.

Hodgman, T.C.: A new superfamily of replicative proteins. Nature 333 (1988) 22-23.

Homerova, D., Sevcikova, B., Sprusansky, O. and Kormanec, J.: Identification of DNA-binding proteins involved in regulation of expression of the *Streptomyces aureofaciens sigF* gene, which encodes sporulation sigma factor σ^F. Microbiology 146 (2000) 2919-28.

Hopwood, D.A.: Forty years of genetics with *Streptomyces* from *in vivo* through *in vitro* to *in silico*. Microbiology 145 (1999) 2183-2202.

Hopwood, D.A., Bibb, M.J., Chater, K.F., Kieser, T., Bruton, C.J., Kieser, H.M., Lydiate, D.J., Smith, C.P., Ward, J.M. and Schrempf, H.: Genetic manipulation of *Streptomyces*: A Laboratory Manual. John Innes Foundation, Norwich, UK., 1985.

Huang, J., Lih, C.J., Pan, K.H. and Cohen, S.N.: Global analysis of growth phase responsive gene expression and regulation of antibiotic biosynthetic pathways in *Streptomyces coelicolor* using DNA microarrays. Genes Dev 15 (2001) 3183-92.

Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. and Omura, S.: Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. Nat Biotechnol 21 (2003) 526-31.

Jankowsky, E., Gross, C.H., Shuman, S. and Pyle, A.M.: The DExH protein NPH-II is a processive and directional motor for unwinding RNA. Nature 403 (2000) 447-51.

Jankowsky, E., Gross, C.H., Shuman, S. and Pyle, A.M.: Active disruption of an RNA-protein interaction by a DExH/D RNA helicase. Science 291 (2001) 121-5.

Jin, L. and Peterson, D.L.: Expression, isolation, and characterization of the hepatitis C virus ATPase/RNA helicase. Arch Biochem Biophys 323 (1995) 47-53.

Jones, P.G., Mitta, M., Kim, Y., Jiang, W. and Inouye, M.: Cold shock induces a major ribosomal-associated protein that unwinds double-stranded RNA in *Escherichia coli*. Proc Natl Acad Sci U S A 93 (1996) 76-80.

Kang, J.G., Paget, M.S., Seok, Y.J., Hahn, M.Y., Bae, J.B., Hahn, J.S., Kleanthous, C., Buttner, M.J. and Roe, J.H.: RsrA, an anti-sigma factor regulated by redox change. Embo J 18 (1999) 4292-8.

Kang, L.W., Cho, H.S., Cha, S.S., Chung, K.M., Back, S.H., Jang, S.K. and Oh, B.H.: Crystallization and preliminary X-ray crystallographic analysis of the helicase domain of hepatitis C virus NS3 protein. Acta Crystallogr D Biol Crystallogr 54 (Pt 1) (1998) 121-3.

Karandikar, A., Sharples, G.P. and Hobbs, G.: Differentiation of *Streptomyces coelicolor* A(3)2 under nitrate-limited conditions. Microbiology 143 (1997) 3581-90.

Keijser, B.J., van Wezel, G.P., Canters, G.W., Kieser, T. and Vijgenboom, E.: The *ram*-dependence of *Streptomyces lividans* differentiation is bypassed by copper. J Mol Microbiol Biotechnol 2 (2000) 565-74.

Keijser, B.J., van Wezel, G.P., Canters, G.W. and Vijgenboom, E.: Developmental regulation of the *Streptomyces lividans ram* genes: involvement of RamR in regulation of the *ramCSAB* operon. J Bacteriol 184 (2002) 4420-9.

Kelemen, G.H., Brian, P., Flardh, K., Chamberlin, L., Chater, K.F. and Buttner, M.J.: Developmental regulation of transcription of *whiE*, a locus specifying the polyketide spore pigment in *Streptomyces coelicolor* A3 (2). J Bacteriol 180 (1998) 2515-21.

Kelemen, G.H., Brown, G.L., Kormanec, J., Potuckova, L., Chater, K.F. and Buttner, M.J.: The positions of the sigma-factor genes, *whiG* and *sigF*, in the hierarchy controlling the development of spore chains in the aerial hyphae of *Streptomyces coelicolor* A3(2). Mol Microbiol 21 (1996) 593-603.

Kelemen, G.H. and Buttner, M.J.: Initiation of aerial mycelium formation in *Streptomyces*. Curr Opin Microbiol 1 (1998) 656-62.

Kelemen, G.H., Viollier, P.H., Tenor, J., Marri, L., Buttner, M.J. and Thompson, C.J.: A connection between stress and development in the multicellular prokaryote *Streptomyces coelicolor* A3(2). Mol Microbiol 40 (2001) 804-14.

Khokhlov, A.S., Tovarova, I.I., Borisova, L.N., Pliner, S.A., Schevchenko, L.A. and Kornitskaya, E.Y.: A-factor responsible for the biosynthesis of streptomycin by a mutant strain of *Actinomyces streptomycini*. Dokl Akad Nauk SSSR 177 (1967) 232-35.

Kim, D.W., Gwack, Y., Han, J.H. and Choe, J.: Towards defining a minimal functional domain for NTPase and RNA helicase activities of the hepatitis C virus NS3 protein. Virus Res 49 (1997a) 17-25.

Kim, D.W., Kim, J., Gwack, Y., Han, J.H. and Choe, J.: Mutational analysis of the hepatitis C virus RNA helicase. J Virol 71 (1997b) 9400-9.

Kirby, K.S., Fox-Carter, E. and Guest, M.: Isolation of deoxyribonucleic acid ribosomal ribonucleic acid from bacteria. Biochemistry Journal 105 (1967) 258-262.

Kodani, S., Hudson, M.E., Durrant, M.C., Buttner, M.J., Nodwell, J.R. and Willey, J.M.: The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene ramS in Streptomyces coelicolor. Proc Natl Acad Sci U S A 101 (2004) 11448-53.

Kormanec, J., Sevcikova, B., Halgasova, N., Knirschova, R. and Rezuchova, B.: Identification and transcriptional characterization of the gene encoding the stress-response sigma factor σ^H in *Streptomyces coelicolor* A3(2). FEMS Microbiol Lett 189 (2000) 31-8.

Korn-Wendisch, F. and Kutzner, H.: A handbook on the biology of bacteria: ecophysiology, isolation, identification, applications, 2 ed. Springer-Verlag, 1992.

Korolev, S., Yao, N., Lohman, T.M., Weber, P.C. and Waksman, G.: Comparisons between the structures of HCV and Rep helicases reveal structural similarities betwen SF1 and SF2 superfamilies of helicases. Protein Sci. 7 (1998).

Kossen, K., Karginov, F.V. and Uhlenbeck, O.C.: The carboxy-terminal domain of the DExDH protein YxiN is sufficient to confer specificity for 23S rRNA. J Mol Biol 324 (2002) 625-36.

Kossen, K. and Uhlenbeck, O.C.: Cloning and biochemical characterization of *Bacillus subtilis* YxiN, a DEAD protein specifically activated by 23S rRNA: delineation of a novel sub-family of bacterial DEAD proteins. Nucleic Acids Res 27 (1999) 3811-20.

Kressler, D., Linder, P. and de La Cruz, J.: Protein trans-acting factors involved in ribosome biogenesis in Saccharomyces cerevisiae. Mol Cell Biol 19 (1999) 7897-912.

Kroos, L. and Maddock, J.R.: Prokaryotic Development: Emerging Insights. J Bacteriol 185 (2003) 1128-46.

Kujat-Choy, S.L.: A redox-regulated RNA helicase gene, Biological Sciences. University of Alberta, Edmonton, 2000, pp. 230. PhD Thesis

Kurdistani, S.K. and Grunstein, M.: *In vivo* protein-protein and protein-DNA crosslinking for genomewide binding microarray. Methods 31 (2003) 90-95.

Kwak, J., McCue, L.A. and Kendrick, K.E.: Identification of bldA mutants of Streptomyces griseus. Gene 171 (1996) 75-8.

Lasko, P.F. and Ashburner, M.: The product of the *Drosophila* gene vasa is very similar to eukaryotic initiation factor-4A. Nature 335 (1988) 611-7.

Lawlor, E.J., Baylis, H.A. and Chater, K.F.: Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). Genes Dev 1 (1987) 1305-10.

Leskiw, B.K., Bibb, M.J. and Chater, K.F.: The use of a rare codon specifically during development? Mol Microbiol 5 (1991a) 2861-7.

Leskiw, B.K., Lawlor, E.J., Fernandez-Abalos, J.M. and Chater, K.F.: TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. Proc Natl Acad Sci U S A 88 (1991b) 2461-5.

Lezhava, A., Kameoka, D., Sugino, H., Goshi, K., Shinkawa, H., Nimi, O., Horinouchi, S., Beppu, T. and Kinashi, H.: Chromosomal deletions in *Streptomyces griseus* that remove the afsA locus. Mol Gen Genet 253 (1997) 478-83.

Lin, C. and Kim, J.L.: Structure-based mutagenesis study of hepatitis C virus NS3 helicase. J Virol 73 (1999) 8798-807.

Linder, P., Lasko, P.F., Leroy, P., Nielsen, P.J., Nishi, K., Schnier, J. and Slonimski, P.P.: Birth of the D-E-A-D box. Nature 337 (1989) 121-22.

Locatelli, G.A., Spadari, S. and Maga, G.: Hepatitis C Virus NS3 ATPase/Helicase: An ATP switch regulates the cooperativity among the different substrate binding sites. Biochemistry 41 (2002) 10332-10342.

Lorsch, J.R.: RNA chaperones exist and DEAD box proteins get a life. Cell 109 (2002) 797-800.

Ma, H. and Kendall, K.: Cloning and analysis of a gene cluster from *Streptomyces coelicolor* that causes accelerated aerial mycelium formation in *Streptomyces lividans*. J Bacteriol 176 (1994) 3800-11.

Merrick, M.J.: A morphological and genetic mapping study of bald colony mutants of *Streptomyces coelicolor*. J Gen Microbiol 96 (1976) 299-315.

Miczak, A., Kaberdin, V.R., Wei, C.L. and Lin-Chao, S.: Proteins associated with RNase E in a multicomponent ribonucleolytic complex. Proc Natl Acad Sci U S A 93 (1996) 3865-9.

Miguelez, E.M., Hardisson, C. and Manzanal, M.B.: Hyphal death during colony development in *Streptomyces antibioticus*: morphological evidence for the existence of a process of cell deletion in a multicellular prokaryote. J Cell Biol 145 (1999) 515-25.

Miguelez, E.M., Hardisson, C. and Manzanal, M.B.: Streptomycetes: a new model to study cell death. Inter Microb 3 (2000) 153-58.

Mohr, S., Stryker, J.M. and Lambowitz, A.M.: A DEAD-box protein functions as an ATP-dependent RNA chaperone in group I intron splicing. Cell 109 (2002) 769-79.

Molle, V. and Buttner, M.J.: Different alleles of the response regulator gene *bldM* arrest *Streptomyces coelicolor* development at distinct stages. Mol Microbiol 36 (2000) 1265-78.

Molle, V., Palframan, W.J., Findlay, K.C. and Buttner, M.J.: WhiD and WhiB, homologous proteins required for different stages of sporulation in *Streptomyces coelicolor* A3(2). J Bacteriol 182 (2000) 1286-95.

Nguyen, K.T., Willey, J.M., Nguyen, L.D., Nguyen, L.T., Viollier, P.H. and Thompson, C.J.: A central regulator of morphological differentiation in the multicellular bacterium *Streptomyces coelicolor*. Mol Microbiol 46 (2002) 1223-38.

Nicol, S.M. and Fuller-Pace, F.V.: The "DEAD box" protein DbpA interacts specifically with the peptidyltransferase center in 23 S rRNA. Proc Natl Acad Sci U S A 92 (1995) 11681-85.

Nodwell, J.R. and Losick, R.: Purification of an extracellular signaling molecule involved in production of aerial mycelium by *Streptomyces coelicolor*. J Bacteriol 180 (1998) 1334-7.

Nodwell, J.R., McGovern, K. and Losick, R.: An oligopeptide permease responsible for the import of an extracellular signal governing aerial mycelium formation in *Streptomyces coelicolor*. Mol Microbiol 22 (1996) 881-93.

Nodwell, J.R., Yang, M., Kuo, D. and Losick, R.: Extracellular complementation and the identification of additional genes involved in aerial mycelium formation in *Streptomyces coelicolor*. Genetics 151 (1999) 569-84.

O'Connor, T.J., Kanellis, P. and Nodwell, J.R.: The *ramC* gene is required for morphogenesis in *Streptomyces coelicolor* and expressed in a cell type-specific manner under the direct control of RamR. Mol Microbiol 45 (2002) 45-57.

Onaka, H., Nakagawa, T. and Horinouchi, S.: Involvement of two A-factor receptor homologues in *Streptomyces coelicolor* A3(2) in the regulation of secondary metabolism and morphogenesis. Mol Microbiol 28 (1998) 743-53.

Paget, M.S., Kang, J.G., Roe, J.H. and Buttner, M.J.: σ^R, an RNA polymerase sigma factor that modulates expression of the thioredoxin system in response to oxidative stress in *Streptomyces coelicolor* A3(2). Embo J 17 (1998) 5776-82.

Pang, P.S., Jankowsky, E., Planet, P.J. and Pyle, A.M.: The hepatitis C viral NS3 protein is a processive DNa helicase with cofactor enhanced RNA unwinding. Embo J 21 (2002) 1168-76.

Pause, A., Methot, N. and Sonenberg, N.: The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis. Mol Cell Biol 13 (1993) 6789-98.

Pause, A. and Sonenberg, N.: Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. Embo J 11 (1992) 2643-54.

Plaskitt, K.A. and Chater, K.F.: Influences of developmental genes on localized glycogen deposition in colonies of a mycelial prokaryote, *Streptomyces coelicolor* A3(2): a possible interface between metabolism and morphogenesis. Phil Trans R Soc B 347 (1995) 105-21.

Pope, M.K., Green, B. and Westpheling, J.: The *bldB* gene encodes a small protein required for morphogenesis, antibiotic production, and catabolite control in *Streptomyces coelicolor*. J Bacteriol 180 (1998) 1556-62.

Pope, M.K., Green, B.D. and Westpheling, J.: The *bld* mutants of *Streptomyces coelicolor* are defective in the regulation of carbon utilization, morphogenesis and cell-cell signalling. Mol Microbiol 19 (1996) 747-56.

Potuckova, L., Kelemen, G.H., Findlay, K.C., Lonetto, M.A., Buttner, M.J. and Kormanec, J.: A new RNA polymerase sigma factor, σ^F, is required for the late stages of morphological differentiation in *Streptomyces* spp. Mol Microbiol 17 (1995) 37-48.

Py, B., Higgins, C.F., Krisch, H.M. and Carpousis, A.J.: A DEAD-box RNA helicase in the Escherichia coli RNA degradosome. Nature 381 (1996) 169-72.

Regnier, P. and Arraiano, C.M.: Degradation of mRNA in bacteria: emergence of ubiquitous features. BioEssays 22 (2000) 235-44.

Rocak, S. and Linder, P.: DEAD-BOX Proteins: The driving forces behind RNA metabolism. Nature 5 (2004) 232-240.

Rogers, G.J., Richter, N. and Merrick, W.: Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. J Biol Chem 274 (1999) 12236-12244.

Rossler, O.G., Straka, A. and Stahl, H.: Rearrangement of structured RNA via branch migration structures catalysed by the highly related DEAD-box proteins p68 and p72. Nucleic Acids Res 29 (2001) 2088-96.

Rozen, F., Edery, I., Meerovitch, K., Dever, T.E., Merrick, W.C. and Sonenberg, N.: Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. Mol Cell Biol 10 (1990) 1134-44.

Ryding, N.J., Bibb, M.J., Molle, V., Findlay, K.C., Chater, K.F. and Buttner, M.J.: New sporulation loci in *Streptomyces coelicolor* A3(2). J Bacteriol 181 (1999) 5419-25.

Ryding, N.J., Kelemen, G.H., Whatling, C.A., Flardh, K., Buttner, M.J. and Chater, K.F.: A developmentally regulated gene encoding a repressor-like protein is essential for sporulation in *Streptomyces coelicolor* A3(2). Mol Microbiol 29 (1998) 343-57.

Sambrook, J., Fritsch, E.F. and Maniatis, T.: Molecular cloning. A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY., 1989.

Sanger, F., S. Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74 (1977) 5463-5467.

Schmid, S.R. and Linder, P.: D-E-A-D protein family of putative RNA helicases. Mol Microbiol 6 (1992) 283-91.

Schwer, B.: A new twist on RNA helicases: DExH/D box proteins as RNPases. Nat Struct Biol 8 (2001) 113-6.

Schwer, B. and Gross, C.H.: Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. Embo J 17 (1998) 2086-94.

Schwer, B. and Meszaros, T.: RNA helicase dynamics in pre-mRNA splicing. Embo J 19 (2000) 6582-91.

Sevcikova, B., Benada, O., Kofronova, O. and Kormanec, J.: Stress-response sigma factor σ^H is essential for morphological differentiation of *Streptomyces coelicolor* A3(2). Arch Microbiol 177 (2001) 98-106.

Shuman, S.: Vaccinia virus RNA helicase: an essential enzyme related to the DE-H family of RNA-dependent NTPases. Proc Natl Acad Sci U S A 89 (1992) 10935-9.

Soliveri, J.A., Gomez, J., Bishai, W.R. and Chater, K.F.: Multiple paralogous genes related to the *Streptomyces coelicolor* developmental regulatory gene *whiB* are present in *Streptomyces* and other actinomycetes. Microbiology 146 (2000) 333-43.

Southern, E.M.: Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology 98 (1975) 503-517.

Stoehr, J.J.: A putative RNA helicase in *Streptomyces coelicolor*, Biological Sciences. University of Alberta, Edmonton, 2001, pp. 191.

Story, R.M., Li, H. and Abelson, J.N.: Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. Proc Natl Acad Sci U S A 98 (2001) 1465-70.

Susstrunk, U., Pidoux, J., Taubert, S., Ullmann, A. and Thompson, C.J.: Pleiotropic effects of cAMP on germination, antibiotic biosynthesis and morphological development in *Streptomyces coelicolor*. Mol Microbiol 30 (1998) 33-46.

Takano, E., Chakraburtty, R., Nihira, T., Yamada, Y. and Bibb, M.J.: A complex role for the γ-butyrolactone SCB1 in regulating antibiotic production in *Streptomyces coelicolor* A3(2). Mol Microbiol 41 (2001) 1015-28.

Takano, E., Nihira, T., Hara, Y., Jones, J.J., Gershater, C.J., Yamada, Y. and Bibb, M.: Purification and structural determination of SCB1, a γ-butyrolactone that elicits antibiotic production in *Streptomyces coelicolor* A3(2). J Biol Chem 275 (2000) 11010-6.

Tan, H. and Chater, K.F.: Two developmentally controlled promoters of *Streptomyces coelicolor* A3(2) that resemble the major class of motility-related promoters in other bacteria. J Bacteriol 175 (1993) 933-40.

Tan, H., Yang, H., Tian, Y., Wu, W., Whatling, C.A., Chamberlin, L.C., Buttner, M.J., Nodwell, J. and Chater, K.F.: The *Streptomyces coelicolor* sporulation-specific

sigma WhiG form of RNA polymerase transcribes a gene encoding a ProX-like protein that is dispensable for sporulation. Gene 212 (1998) 137-46.

Tanner, N. and Linder, P.: DExD/H box RNA helicases: from generic motors to specific dissociation functions. Molecular Cell 8 (2001a) 251-262.

Tanner, N.K. and Linder, P.: DExD/H box RNA helicases: from generic motors to specific dissociation functions. Mol Cell 8 (2001b) 251-62.

Thompson, C.J., Fink, D. and Nguyen, L.D.: principles of microbial alchemy: insights from the *Streptomyces coelicolor* genome sequence. Genome Biol 3 (2002) 1020.1-1020.4.

Tillotson, R.D., Wosten, H.A., Richter, M. and Willey, J.M.: A surface active protein involved in aerial hyphae formation in the filamentous fungus *Schizophillum commune* restores the capacity of a bald mutant of the filamentous bacterium *Streptomyces coelicolor* to erect aerial structures. Mol Microbiol 30 (1998) 595-602.

Tseng, S., Weaver, P., Liu, Y., Hitomi, M., Tartakoff, A. and Chang, T.: Dbp5p, a cytosolic RNa helicase, is required for poly (A)+RNA export. Embo J 17 (1998) 2651-2662.

Velankar, S.S., Sultanas, P., Dillingham, M.S., Subramanya, H.S. and Wigley, D.B.: Crystal structures of complexes of PcrA DNA helicase with a DNA usbstrate indicate an inchworm mechanism. Cell 97 (1999) 75-84.

Viollier, P.H., Kelemen, G.H., Dale, G.E., Nguyen, K.T., Buttner, M.J. and Thompson, C.J.: Specialized osmotic stress response systems involve multiple SigB-like sigma factors in *Streptomyces coelicolor*. Mol Microbiol 47 (2003) 699-714.

Viollier, P.H., Minas, W., Dale, G.E., Folcher, M. and Thompson, C.J.: Role of acid metabolism in *Streptomyces coelicolor* morphological differentiation and antibiotic biosynthesis. J Bacteriol 183 (2001a) 3184-92.

Viollier, P.H., Nguyen, K.T., Minas, W., Folcher, M., Dale, G.E. and Thompson, C.J.: Roles of aconitase in growth, metabolism, and morphological differentiation of *Streptomyces coelicolor*. J Bacteriol 183 (2001b) 3193-203.

Vohradsky, J., Li, X.M., Dale, G., Folcher, M., Nguyen, L., Viollier, P.H. and Thompson, C.J.: Developmental control of stress stimulons in *Streptomyces coelicolor* revealed by statistical analyses of global gene expression patterns. J Bacteriol 182 (2000) 4979-86.

von Hippel, P.H. and Delagoutte, E.: A general model for nucleic acid helicases and their "coupling" within macromolecular machines. Cell 104 (2001) 177-90.

von Hippel, P.H. and Delagoutte, E.: Macromolecular complexes that unwind nucleic acids. BioEssays 25 (2003) 1168-1177.

Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J.: Distantly related sequences in the alph- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. Embo J 1 (1982) 945-51.

Watty, A., Methfessel, C. and Hucho, F.: Fixation of allosteric states of the nicotinic acetylcholine receptor by chemical cross-linking. Proc Natl Acad Sci U S A 94 (1997) 8202-8207.

Watve, M.G., Tickoo, R., Jog, M.M. and Bhole, B.D.: How many antibiotics are produced by the genus *Streptomyces*? Arch Microbiol 176 (2001) 386-390.

White, J. and Bibb, M.: *bldA* dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. J Bacteriol 179 (1997) 627-33.

Wildermuth, H.: Development and organization of the aerial mycelium in *Streptomyces coelicolor*. J Gen Microbiol 60 (1970) 43-50.

Wildermuth, H. and Hopwood, D.A.: Septation during sporulation in *Streptomyces coelicolor*. J Gen Microbiol 60 (1970) 51-59.

Willey, J., Santamaria, R., Guijarro, J., Geistlich, M. and Losick, R.: Extracellular complementation of a developmental mutation implicates a small sporulation protein in aerial mycelium formation by *S. coelicolor*. Cell 65 (1991) 641-50.

Willey, J., Schwedock, J. and Losick, R.: Multiple extracellular signals govern the production of a morphogenetic protein involved in aerial mycelium formation by *Streptomyces coelicolor*. Genes Dev 7 (1993) 895-903.

Yao, N., Hesson, T., Cable, M., Hong, Z., Kwong, A.D., Le, H.V. and Weber, P.C.: Structure of the hepatitis C virus RNA helicase domain. Nat Struct Biol 4 (1997) 463-7.

Yu, E. and Owttrim, G.: Characterization of the cold stress-induced cyanobacterial DEAD-box protein CrhC as an RNA helicase. Nucleic Acids Res 28 (2000) 3926-34.

Zhen, L. and Swank, R.T.: A simple and high yield method for recovering DNA from agarose gels. BioTechniques 14 (1993) 894-898.